THE ROLE OF SRY-BOX (SOX) TRANSCRIPTION FACTORS IN EPITHELIAL STEM CELL BIOLOGY OF THE GASTROINTESTINAL TRACT

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 Master of Science in the Department of Cell and Molecular Physiology.

Chapel Hill
2011

Approved by:

Scott T. Magness, Ph.D.

P. Kay Lund, Ph. D.

Scott Randell, Ph.D.

ABSTRACT

ADAM DAVID GRACZ: THE ROLE OF SRY-BOX (SOX) TRANSCRIPTION FACTORS IN EPITHELIAL STEM CELL BIOLOGY OF THE GASTROINTESTINAL TRACT

(Under the direction of: Scott T. Magness, Ph.D. and P. Kay Lund, Ph.D.)

Stem cell biology, though a well-established concept in the scientific zeitgeist, is only beginning to emerge as an independent field of study. An understanding of basic stem cell biology and its translation into clinically applicable therapies holds the potential to direct a paradigm shift in modern medical practice. However, in order to realize this potential, the scientific community must first understand the genetic and molecular mechanisms by which cells acquire and maintain 'stemness', specifically, multipotency and the ability to self-renew. Studying the transcriptional machinery that controls these properties could lead to a greater understanding of stem cells as a component of physiology as a whole. Due to its critical importance to homeostasis, the gastrointestinal tract is an attractive system for the study of stem cell biology. *Sox* factors, a group of transcription factors that have previously been associated with embryonic and neural stem cells, are rapidly emerging as central to maintaining 'stemness' in the gastrointestinal tract as well. This work reviews the known role of *Sox* factors in the gastrointestinal epithelium and describes our novel findings regarding *Sox9* as a marker of stem cells in the adult intestinal epithelium.

To my mentors for their continued support, guidance, and encouragement.

ACKNOWLEDGMENTS

The work presented in this thesis would not have been possible without the help, guidance, and support of a number of individuals. I would like to first thank my mentors, Drs. Scott Magness, P Kay Lund, and Duane Mitchell, 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, and Victoria Bali. Additionally, our group has been fortunate enough to enlist the support of a group of extraordinarily talented undergraduates. I would like to acknowledge Sendhinathan Ramalingam, Kyle Roche, Michael Cronce, and Kelly Lindblom for their support of the work presented in this thesis.

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 thesis.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

\

LIST OF ABBREVIATIONS

Alfp	alpha-fetoprotein
APC	allophycocyanin
CBC	crypt base columnar
CD	clusters of differentiation
cDNA	complementary deoxyribonucleic acid
CEA	carcinoembryonic antigen
DNA	deoxyribonucleic acid
DTT	dithiotriotol
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
EGFP	enhanced green fluorescent protein
EpCAM	epithelial cell adhesion molecule
FACS	fluorescent activated cell sorting
FBS	fetal bovine serum
GENSAT	gene expression nervous system atlas
GI	gastrointestinal
HBSS	Hank's balanced salt solution
HESC	human embryonic stem cell
HMG	high mobility group
hr	hour
IBD	inflammatory bowel disease
IESC	intestinal epithelial stem cell

IM	intestinal metaplasia
Lgr5	leucine-rich repeat-containing G-protein coupled receptor 5
min	minute
mRNA	messenger ribonucleic acid
NDS	normal donkey serum
NGS	normal goat serum
OCT	optimal cutting temperature
PanIN	pancreatic intraepithelial neoplasia
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PFA	paraformaldehyde
POU domain	Pit-1 Oct1/2 Unc-86 domain
PP	pancreatic polypeptide
RNA	ribonucleic acid
RPM	revolutions per minute
RT-PCR	reverse transcriptase polymerase chain reaction
sec	second
SMA	smooth muscle actin
Sox	*Sry* Box-containing
Sry	Sex determining region Y
TDF	testis determining factor
TEF	tracheoesophageal fistula
TGF-β	transforming growth factor β

CHAPTER 1

SCOPE OF WORK

The work presented in this thesis consists largely of two peer-reviewed and published manuscripts. The first, presented in Chapter 2, is a comprehensive review of the *Sox* family of transcription factors and their role in the epithelial tissues of the gastrointestinal tract. Chapter 3 focuses on original research into specific properties of *Sox9* in the intestinal epithelium.

We chose to pursue investigation of *Sox* factors in the gastrointestinal tract due to their generally accepted function as ubiquitous components of stem cell specific molecular signaling, as well as their critical role in maintaining stemness and regulating differentiation. Our work identified and validated that *Sox9* is an intestinal epithelial stem cell (IESC) biomarker. In addition to this major finding, the work presented here establishes critical techniques and reagents for the isolation and *in vitro* study of IESCs, and lays the groundwork for mechanistic study of *Sox* factors in IESCs and progenitors.

CHAPTER 2

SRY-BOX (SOX) TRANSCRIPTION FACTORS IN GASTROINTESTINAL PHYSIOLOGY AND DISEASE

ADAM D GRACZ AND SCOTT T MAGNESS

INTRODUCTION

Contemporary biomolecular research promises to advance gastroenterology in the 21st century by providing clinicians and researchers with a growing understanding of the molecular mechanisms that underlie both normal physiology and disease of the gastrointestinal tract. As the basic and clinical research communities grow closer to bridging the gap between bench and bedside, it is apparent that an understanding of cellular maintenance and tissue repair mechanisms is central to achieving this goal. *Sox* factors are a family of structurally related transcription factors that are emerging as regulators of transcriptional activity with potent effects on cellular phenotypes. This review will focus on the current understanding of the role of *Sox* factors in normal cellular maintenance and differentiation, as well as in disease states of the endodermally derived tissues of the gastrointestinal tract.

The gastrointestinal tract, which includes the esophagus, stomach, and small and large intestine, is constantly exposed to microbes, chemical toxins or mutagens, varying pH, and physical injury to the epithelial barrier. The consequences of compromising the epithelial barrier of the gastrointestinal tract are leakage of bacteria into the surrounding vasculature, leading to sepsis; or leakage of acid into the underlying mesenchyme resulting in ulcers and chronic inflammation. Therefore, tissue renewal driven by

epithelial stem cells is critical to maintaining constant integrity of epithelial barriers and ultimately the survival of the organism.

In addition to maintaining basal function, tissue-specific stem cells are capable of responding to injury, damage, or even large-scale loss of tissue to attempt repair. For example, intestinal epithelial stem cells undergo expansion and drive massive tissue remodeling following ileo-cecal resection (32). However, stem cell populations do not always respond equally to damage. Following pancreatectomy, ductal cell populations of the rat pancreas are able to regenerate a large portion of the organ's original mass, but in the case of type I diabetes, there are no regenerative processes that replace destroyed β-islet cells (16). Further complicating the understanding of stem cell mediated regeneration is the observation that conventional animal models do not always exhibit responses that translate to human biology. For instance, recent findings indicate that β-islet cell regeneration does not occur following pancreatectomy in adult human patients, as previously observed in rats (16, 86). Aberrant proliferation of tissue-specific stem cells is implicated in a wide range of common diseases throughout the gastrointestinal tract, including diabetes, cirrhosis, and gastrointestinal cancers (82, 84, 117, 159). The differential abilities of stem cell pools to respond to regenerative stimuli point to intrinsic genetic components that influence the stemness of the cell and are capable of reacting to extrinsic cues.

Attempts to expand on the nascent understanding of stem cell maintenance and differentiation in the gastrointestinal tract has led to the identification of several gene families implicated in the growth and expansion of multipotent stem cell populations. Among these families is a group of genes known as *Sox* (*Sry* box-containing) transcription factors. Founding members of the *Sox* family were first described in terms

of their role in establishing sexual dimorphism in development (46, 70, 118). Shortly thereafter, *Sox* factors were identified in adult neural cell populations where they were shown to have powerful roles in maintaining tissue-specific stem cell populations within the nervous system (27, 118, 137).

A common theme in recent studies is that *Sox*-factors possess a potent capacity to direct or influence 'stemness', or a cell's ability to meet the established stem cell criteria of multipotency and self-renewal. Landmark studies demonstrated that *Sox2*, along with three other transcription factors, possessed the ability to reprogram differentiated adult cells to a state of pluripotency resembling that seen in embryonic stem cells (129, 130). Taken with the cell-specific expression patterns and redundant function of *Sox* factors, the demonstrated ability of *Sox2* to regulate cellular potency suggests that *Sox* factors might play a role in maintaining 'stemness' in cells of the gastrointestinal tract.

SOX FACTORS: VERSATILE REGULATORS OF TRANSCRIPTIONAL ACTIVITY

There are currently thirty described *Sox* factors in mammals. The first *Sox* gene characterized was *Sry* (sex-determining region Y). *Sry* was identified as the previously reported testis-determining factor, *TDF*, in sex-reversed human XX males and XY females (118). *Sox* factors represent a family of genes within the high mobility group (HMG) superfamily, demonstrating homology in their 79 amino acid HMG-box DNA-binding domain motifs (46, 118, 122). All *Sox* factors exhibit a conserved HMG domain motif of RFMNAF, which distinguishes them from other HMG-box genes (14). However, it has been observed that all *Sox* factors with the exception of *Sry* have an extended motif homology of RFMNAFMVW and it has been suggested that this

homology sequence be used for classification of *Sox* factors (14). The HMG-box has the capacity to bind both DNA consensus elements and other transcriptional regulators, such as POU-domain proteins, to modulate transcriptional activity (5, 157). All *Sox* factors studied to date demonstrate the ability to bind a DNA consensus sequence of (A/T)(A/T)CAA(A/T)G, allowing these factors to sometimes exhibit functionally redundant roles (27, 28, 34, 42, 50, 53, 138). Phylogenic analysis of *Sox* factor HMG-box sequence and protein structure has identified similarities between different *Sox* genes that has resulted in the classification of *Sox* factors into subgroups A-J (14).

Sox factors share a striking ability to affect structural changes in DNA. *Sox* factors have been shown to bind the minor groove of the DNA helix, resulting in a significant degree of DNA bending not elicited by other members of the HMG superfamily that bind to specific DNA sequences (37, 42). The participation of *Sox* factors in regulating a conformational change in DNA suggests that these factors may possess powerful transcriptional modulating roles that go beyond the function of site-specific transcriptional activation and implies an effect on larger-order chromatin structure (37).

Observations regarding the architectural effect of *Sox* factors on DNA highlight the significance of this property. Early studies on the role of *Sry* in sex reversal revealed that most aberrant mutations occur in the HMG domain, suggesting that DNA binding and bending properties are essential in sex determination (37, 97). It has been proposed that by affecting a dramatic conformational change in DNA, *Sox* factors are capable of bridging the gap between distal enhancers and proximal transcriptional elements, allowing for interaction between regulatory complexes separated by long distances in the genome (94). Furthermore, it has been demonstrated that the *Sox* HMG box has an adaptive tertiary structure, the formation of which is directed by DNA upon binding

consensus sequences (29, 142). The ability of the *Sox* HMG box to change tertiary structure depending on DNA-binding has led to the proposal that single *Sox* factors are capable of bending DNA to different degrees in a context-dependent manner, facilitating differential transcriptional control of genes (147).

Functional roles of Sox factors are dependent on spatial, dose, and temporal context

Sox factors are expressed in nearly every tissue during embryogenesis and are emerging as a group of genes that are broadly expressed across a wide range of adult tissues as well. Individual *Sox* factors are expressed across many tissue and proliferating cell types, suggesting fundamental roles in cell maintenance. However, *Sox* factors exhibit some degree of cell- and tissue-specificity, in that they are expressed in multiple tissues but not in every cell found in those tissues. For example, *Sox9* is expressed in the Sertoli cells of the testis, in chondrocytes during cartilage formation, and in a subset of epithelial cells of the small intestine and colon (39, 51, 143, 153). This characteristic of expression patterns that are not restricted to one organ systems or all cell types within tissues arising from the same germ layer suggests that *Sox* factors have a fundamental functional role in cell behavior across diverse tissue and cell types.

The ability to bind the same DNA sequence is predictive of redundancy between different *Sox* factors and the potential to functionally compensate for one another (37, 70). Compensatory mechanisms have been suggested by the observation that $Sox5^{null}$ and $Sox6^{null}$ mice each undergo chondrogenesis, but animals with deletions of both *Sox5* and *Sox6* fail to produce cartilage (120). These data imply that *Sox5* and *Sox6* are able to reciprocally compensate for one another and drive the development of cartilage.

Another important feature of *Sox* factors is that they can act in a dose-dependent manner. That is, varying levels of expression of a single *Sox* factor can elicit a variety of

responses in distinct cell and tissue types. It has been suggested that gradients of endogenous expression of *Sox*-factors across different cell types in a single tissue allow for different responses to the same exogenous signal (39, 40, 73, 132, 143). For example, data show that differing levels of *Sox2* during development drive retinal tissue to adopt distinct phenotypes (132). Dose-dependent behavior of *Sox* factors was suggested for *Sox9* in human patients suffering from campomelic dysplasia, a disorder defined by significant skeletal defects. It was observed that patients suffering from campomelic dysplasia in addition to sex reversal exhibited a varying range of phenotypic penetrance due to haploinsufficiency (40). Variable phenotypes as opposed to a true wild type vs. "knock-out" phenotypes in campomelic dysplasia suggest that *Sox* factors are able to exert functional effects on gene expression in a dose-dependent manner.

In addition to behaving in a dose-dependent manner, *Sox* factors often exhibit temporal expression that may play an important role in the maintenance and differentiation of cellular lineages. The observed trend has been that the expression of *Sox* factors is maintained prior to cell fate decisions, then rapidly downregulated upon lineage commitment. This has been observed in the development of the testis, during which *Sox9* and *Sry* undergo a change from upregulation to downregulation during precursor commitment to form Sertoli cells, as well as in dynamic expression patterns of *Sox1-3* in neuroepithelial cell lineages (27, 44, 48, 70, 127, 137). Though still in the process of being thoroughly defined, the known transcriptional capabilities of *Sox* factors demonstrate that they are powerful regulators of cellular potency and developmental processes.

Sox17 is essential for gastrointestinal development

The epithelial layers lining the esophagus, stomach, and intestine, as well as a majority of liver and pancreatic tissue, all arise from the definitive endoderm during embryonic development (reviewed by (148). *Sox17*, a member of the SoxF subgroup, was shown to be essential to the development of definitive endoderm in mammals, as had previously been demonstrated for *Sox17* orthologues in *Xenopus* (55, 59). Mice deficient in *Sox17* showed defects in the definitive endoderm beginning with endodermal induction. Specifically, there was an increase in apoptosis observed in the foregut along with defects in cellular expansion in the mid- and hindgut, resulting in embryonic lethality at 10.5 dpc (59). Recent studies have also demonstrated a role for *SOX17* in the differentiation of human embryonic stem cells (HESCs). Overexpression of *SOX17* in HESCs drove 78.51% of the cells to differentiate into definitive endoderm in the absence of cytokines, while treatment with Activin A, a cytokine known to promote formation of definitive endoderm, only drove one-third of HESCs to differentiate similarly (114). Additionally, further manipulation with growth factors associated with lineage-specific differentiation was successful in inducing the expression of early hepatic and pancreatic markers in *SOX17*-overexpressing HESCs (114). Taken together, these findings highlight a critical role for *Sox17* in the development of the mammalian gastrointestinal tract. This established developmental role suggests that *Sox* factors are an important subject of study in the maintenance of stem cell populations in the organs of the adult gastrointestinal tract.

THE UPPER GI TRACT: ESOPHAGUS/STOMACH

Maintenance and differentiation of cells in the upper gastrointestinal tract are important in facilitating the developmental formation and continued function of the normal esophagus and stomach. Due to common embryonic origin of the esophagus and trachea, any aberrant molecular signaling early in development can result in phenotypes that compromise not only the integrity of the digestive tract, but that of the airway and respiratory system as well. In the adult organism, esophageal and gastric epithelial homeostasis is governed by maintenance and differentiation of resident stem and progenitor cells (60). Aberrant proliferation or differentiation in these cell populations can lead to metaplastic pre-cancerous lesions that could ultimately result in functional deficits of the upper digestive tract. Therefore, an understanding of the regulatory factors governing development and maintenance of proper temporal and spatial cellular phenotypes within the upper gastrointestinal tract is essential for the characterization of a wide range of disorders affecting the esophagus and stomach.

Stem cells in the esophagus

The esophageal epithelium is arranged in two layers: the basal layer, consisting of cuboidal epithelial cells, and the suprabasal layer, comprised of polyhedral cells (74). Early research with radioactive thymidine-labeling studies demonstrated that cell division is restricted to the basal layer (74). These studies also demonstrated a relatively high rate of turnover in the esophageal epithelium, as nearly all of the basal epithelial cells were shown to divide in 3 to 5 days. More recent work has gone further to demonstrate that basal esophageal epithelial cells are arranged in clonal units (30, 134). Taken together, these data support the presence of an esophageal stem cell and stem cell niche in the basal layer of the esophageal epithelium. Furthermore, it has been shown that transit-

amplifying progenitor cells generated by stem cells in the basal esophageal epithelium must enter the suprabasal layer before differentiating (30). This observation has led to the proposal that genetic changes must take place within progenitor cells prior to differentiation that allow them to generate appropriate cellular lineages once they enter the suprabasal layer (31). Significant genetic priming for differentiation suggests the involvement of potent transcriptional regulation, a functional requirement that makes *Sox* factors an attractive subject in esophageal stem cell biology.

Stem cells in the stomach

The human gastric epithelium consists of simple columnar epithelium and is divided into the cardia, fundus, and antrum (52). It is important to note that the murine stomach is divided differently, into forestomach lined with stratified squamous epithelium and hindstomach lined with simple columnar epithelium (61). Invaginations, called gastric pits, line the gastric epithelium, each containing 4 or 5 gastric glands, which house the gastric epithelial proliferative zone (61, 75). In both fundic and antral gastric pits, gastric epithelial stem cells reside at the neck of the glands and drive proliferation bi-directionally up and down the gland (60). However, fundic and antral pits differ in terms of cell types generated by their respective stem cells and also in terms of rates of regeneration, with the antral pits driving cellular turnover at a greater rate than the fundic pits (49, 60, 93). This difference in proliferation across separate zones of the same tissue suggests differential genetic regulation of stemness in gastric epithelial stem cell populations. The demonstrated ability of *Sox* factors to regulate proliferative potential and stemness makes them attractive candidates for regulation of proliferation across the gastric epithelium.

Evidence that Sox factors play a role in stem cell physiology of the upper GI tract

During development, *Sox* factors drive tissue specification processes along the axis of the forming upper gastrointestinal tract (102). The esophagus and stomach derive from the endodermal foregut and demonstrate similarities in terms of *Sox* expression. Of the *Sox* factors associated with the upper gastrointestinal tract (Table 1), *Sox17* is expressed in the esophagus and stomach, while *Sox18* expression is reported in the stomach alone (64, 106). Within the normal gastrointestinal tract, *Sox2* shows expression in the esophagus and stomach, and its functional role in the upper gastrointestinal tract has been extensively characterized (77, 102). Interest in the role of *Sox2* in the stomach was originally driven by the observation that *Sox2* expression is markedly downregulated in gastric carcinomas, implicating aberrant expression of the gene with a loss of proper cellular homeostasis (see Table 2) (77). This early observation suggested that *Sox2* might also play a role in maintenance and differentiation processes of gastric stem cell populations.

Sox2 in the development and normal biology of the esophagus and stomach

During early development of the esophagus, *Sox2* is expressed in all cells of the endodermal foregut, but at E9.5 expression levels segregate in a dorsal-ventral manner that results in relatively high expression in the developing esophagus and relatively low expression in the developing trachea (102). Abnormal down-regulation of *Sox2* expression during embryogenesis was observed to result in esophageal atresia (EA) and tracheoesophageal fistula (TEF), with *Sox2* protein undetectable at the site of TEF (92, 102). These data demonstrate that normal expression of *Sox2* is required for appropriate tracheoesophageal morphogenesis.

In the case of murine gastric development, *Sox2* levels are first expressed across the entire embryonic stomach to the duodenum, but decrease in the hindstomach while remaining relatively high in the forestomach later in gestation (102). This gradient in expression levels is mirrored in adult human stomach, with *Sox2* expression decreasing from fundus to antrum and becoming very low at the pyloric sphincter (135). The change in expression of *Sox2* along the developing esophagus and stomach implies a dose-dependent role in the establishment of cellular fate, as high levels of *Sox2* appear to specify an esophageal or gastric fate, while low levels of *Sox2* appear to drive adaptation of tracheal or intestinal fates.

Sox2 plays an important role in cellular differentiation processes that establish the structure of epithelia formed along the length of the upper gastrointestinal tract (102). It has been proposed that high levels of *Sox2* during murine development drive the formation of stratified, squamous epithelium while low levels direct cell fate toward columnar epithelium (92). The change in mucosal structure from stratified squamous epithelium in the murine forestomach to columnar epithelium in the hindstomach supports this concept. Furthermore, mice exhibiting TEF concurrent with decreased levels of *Sox2* expression exhibit ectopic columnar epithelium at the site of TEF, in contrast to wild-type littermates, which have physiologically normal squamous epithelium throughout the esophagus (92). In addition to playing a role in the development of the structural anatomy of the upper gastrointestinal tract, *Sox2* appears to be important in the establishment of proper epithelial subtypes along the esophageal-gastric axis.

Roles for Sox2 in precancerous and cancerous lesions of the esophagus and stomach

Cellular changes in precancerous and cancerous lesions of the esophagus and stomach suggest that *Sox2* may also play an active role in the maintenance and differentiation of cells in adult tissues of the upper GI tract. *Sox2* appears to be important for maintaining a gastric phenotype in epithelial cells in the stomach (102, 131, 135). Failure to maintain gastric phenotype in gastric epithelial cell populations can lead to intestinal metaplasia (IM), or the precancerous conversion of gastric epithelial cells to intestinal epithelial cells (135). Studies of human tissue from patients suffering from IM showed that *Sox2* is progressively down-regulated throughout progression from normal gastric epithelium to widespread intestinal metaplasia (135). A loss of *Sox2* was also observed in both a rat model of intestinal metaplasia and samples from human cases of Barrett's esophagus (9). These observations are strikingly similar to the finding that segregation of tissue expressing *Sox2* from tissue not expressing *Sox2* early in development coincides with the formation of a gastric-duodenal junction (92). Taken together, these studies suggest that *Sox2* is critical in maintaining a distinct boundary between the epithelia of the esophagus and stomach and intestinal epithelia.

Sox2 expression levels have also been shown to associate with other distinct markers of a gastric phenotype. For example, *Sox2* upregulates the expression of the stomach-specific pepsin precursor *pepsinogen A* in gastric cell lines as well as ectopically in colon cancer cell lines (131). Conversely, interference with *Sox2* expression causes a significant decrease in *pepsinogen A*, suggesting a pivotal role for *Sox2* in *pepsinogen A* expression (131). Further evidence for *Sox2* maintenance of a gastric phenotype is found by comparing expression of gastric-specific mucins versus that of the intestinal-specific mucin. When *Sox2* levels decrease, there is a progression from normal gastric phenotype

to intestinal metaplasia, and expression of gastric-specific mucin decreases, while levels of intestinal-specific mucin increase (135). To summarize these studies, high levels of *Sox2* are associated with the upregulation of markers for differentiation specific to the gastric epithelium, while decreasing levels of *Sox2* correlate with the upregulation of a lineage-specific marker normally restricted to the intestinal epithelium.

Gastric cancer is commonly associated with a down-regulation of *Sox2* (77). It has been proposed that methylation and epigenetic silencing of *Sox2* in gastric cancer leads to progressively aberrant proliferation of the gastric epithelium, an idea supported by the clinical observation that patients presenting with gastric carcinomas which exhibit methylation of the *Sox2* gene face poorer prognoses than those without *Sox2* methylation (92). Reinforcing the idea that *Sox2* might play a tumor-suppressive role in the gastric epithelium is the observation that exogenous overexpression of *Sox2* has an inhibitory effect on proliferation of gastric carcinoma cells *in vitro*, an effect that progresses to morphological change and apoptosis with persistent overexpression of *Sox2* (92). Interestingly, a recent study on human esophageal squamous cell carcinoma reported that 15% of patients had amplification of *Sox2* in tissue samples from this tumor type (19). This study raises the possibility that *Sox2* may play a different regulatory role in the esophagus than it does in the stomach. Taken as a whole, the above data suggests that *Sox2* is a critical player in the maintenance of a gastric epithelial cellular phenotype, as well as a regulator of normal proliferation of the esophageal and gastric epithelia.

LOWER GI TRACT

The lower gastrointestinal tract, which includes the small intestine and colon, functions to absorb nutrients in the digestive tract through a complex interaction between

epithelial, vascular, nervous, and lymphatic tissues. In addition, epithelial tissue lining the luminal surfaces of the lower gastrointestinal tract must be specialized for absorptive function while simultaneously maintaining a barrier against the bacteria-rich digestive lumen in order to preserve the sterility of the rest of the body. A majority of epithelial cells carry out highly specialized functions in the intestinal and colonic epithelia and undergo rapid turnover rates that result in replacement of nearly the entire epithelium every 7 days (154). This rate of rapid renewal combined with the need to sustain physiological and cellular homeostasis emphasizes the importance of regulation of cellular maintenance and differentiation in the small intestine and colon. The intestinal epithelium has been an attractive target for the study of cellular maintenance and differentiation, especially in terms of potent mitogenic signaling via the *Wnt* pathway. The role of the intestinal and colonic epithelia in inflammatory bowel disease (IBD) as well as high rates of colon cancer has fueled a long-standing interest in maintenance and differentiation of this tissue type.

Stem cells of the small intestine and colon

The proliferative region of the intestinal and colonic epithelia has long been known to reside in the crypts of Lieberkuhn (22-24). Division of multipotent stem cells and differentiation of progenitor cell populations in the crypts drives the replacement of terminally differentiated cells on the villi. These cells are then sloughed into the lumen facilitating epithelial renewal. In addition to generating absorptive enterocytes that comprise a majority of the terminally differentiated epithelium, the intestinal epithelial stem cells also give rise to secretory lineages that include goblet and enteroendocrine cells in the small intestine and colon, as well as Paneth cells, which are restricted to the small intestine (22-25). Strong evidence points to the *Wnt*/*β-catenin* pathway as a key

mitogenic signaling cascade driving crypt proliferation (26). Significantly less is understood in terms of how *Wnt*/*β-catenin* signaling is regulated in order to elicit appropriate and distinct responses from cells responsible for controlling stem cell maintenance versus differentiation in the small intestine and colonic epithelia. Recent advances in the understanding of *Sox* function in the intestinal epithelium has revealed that this family of transcription factors plays an important role in the downstream regulation of *Wnt* activity (72).

High rates of proliferation, differentiation, and cell turnover make the intestinal epithelium a useful model system to study the role of *Sox* factors in regulating cellular homeostasis in the gastrointestinal tract. Screening for mRNA has revealed expression of *Sox2, Sox3, Sox4, Sox5, Sox6, Sox7, Sox9, Sox10, Sox11, Sox17* and *Sox18* in whole small intestine and colon (see Table 1). A caveat concerning mRNA expression is that some data is reported from whole-tissue analysis. Therefore, some *Sox* factors reported to be expressed in the intestine might be specific to non-epithelial cell types, such as smooth muscle and enteric nerve cells, which will not be addressed in this review. However, the role of *Sox9* in the epithelium of the lower gastrointestinal tract has been extensively characterized (10, 13, 119). Studies utilizing the intestine as a model system have been instrumental in revealing important epithelial roles for *Sox4* and *Sox17* as well (119). Data suggests that *Sox* factors contribute to the proper regulation of the proliferative cell populations in the intestinal and colonic epithelium.

Sox9 regulates proliferation in the small intestine and colon

The most extensively studied *Sox* factor in the intestinal epithelium is *Sox9*, which has been shown to be primarily expressed in the stem/progenitor region of the crypts in the small intestine and colon, with some limited expression observed in cells throughout

the villi of the small intestine (13, 39, 58). The localization of *Sox9* to the proliferative compartment of the intestinal epithelium supported a possible involvement in the *Wnt/β-catenin* pathway, as *Wnt* is known to drive proliferation in the crypts. Disruption of *Wnt* signaling in early post-natal stages was shown to completely ablate *Sox9* expression in the intestinal and colonic epithelia, identifying *Sox9* as a downstream target of *Wnt* (13).

Other studies demonstrate that *Sox9* plays an important role in the regulation of cell fate and differentiation. *Sox9* expression has been shown to prevent cellular differentiation into *Muc2*-expressing goblet cells *in vitro*, implicating *Sox9* in the maintenance of a non-differentiated population of putative stem/progenitor cells (13). Furthermore, *Sox9* was shown to repress genes associated with terminally differentiated cell types, including *CDX2* in intestinal epithelium and carcinoembryonic antigen (*CEA*) in colonic epithelium, *in vitro* (13, 58). More recently, conditional embryonic *Sox9* knockout mice were shown to lack Paneth cells in the intestinal epithelium following birth (10, 89). Taken together, these data suggest that *Sox9* plays a complex regulatory role in cell fate decisions and differentiation downstream of *Wnt*.

Findings suggestive of *Sox9* association with a non-differentiated pool of cells located in the classical "stem cell niche" of the intestinal and colonic epithelia provide evidence that *Sox9* might also play a role in the maintenance of a stem/progenitor phenotype. However, given the multitude of proposed functional roles for *Sox9*, it can become conceptually difficult to grasp how a single gene appropriately regulates different cellular responses across the epithelium. *Sox9* is expressed at variable levels in the intestinal epithelium and these levels have been shown to correlate with different cellular phenotypes, reinforcing earlier observations that *Sox* factors function in dose-dependent manners (39). Specifically, recent studies utilizing a BAC transgenic mouse model, in

which regulatory elements of *Sox9* drive the expression of EGFP, classified *Sox9* expression as “sublow”, “low”, and “high” (39, 45). *Sox9* high-expressing cells were demonstrated to be fully differentiated enteroendocrine cells, as well as the only *Sox9*-positive cell type observed in villi, while *Sox9* low-expressing cells were shown to correlate with increased expression of established functional markers of intestinal epithelial stem cells (39). Recently, it was shown that *Sox9* low-expressing cells are multipotent and self-renewing *in vitro* and that *Sox9* “sublow”-expressing cells are enriched for markers of proliferation as well as early markers of differentiation, suggesting that *Sox9* “sublow” expression corresponds to the transit-amplifying cells of the intestinal epithelium (45). *In vitro* studies have shown that *Sox9* inhibits the transcription of *cMyc* and *cyclinD1* by *β-catenin/TCF*, suggesting that *Sox9* might regulate *Wnt* signaling via a dose-dependent negative feedback loop (10, 13, 39). One model proposes that *Sox9* has a dose-dependent, inverse relationship with proliferative potential in response to *Wnt* signaling in the intestinal crypts (39). This proposal is strongly reinforced by the identification of the *Sox9* “sublow”-expressing putative transit amplifying population, wherein very low levels of *Sox9* would be expected to result in an upregulation of *cMyc* and *cyclinD1* in response to *Wnt* signals, potentiating the rapid proliferation needed to constantly renew the terminally differentiated cells of the intestinal epithelium (45). Interestingly, very high levels of Sox9 are expressed in the enteroendocrine populations, which develop in the absence of *Wnt-signalling* (39, 146). This observation implicates a *Wnt*-independent regulatory role for *Sox9* in post-mitotic cell populations.

Sox9 and colon cancer

In addition to apparent regulatory roles in the normal biology of the small intestine and colon, increased expression of *Sox9* has been noted in colon cancer (see Table 2). Screening of several human colon cancer cell lines revealed that *Sox9* mRNA and protein were upregulated over controls (see Table 3) (13). Immunohistochemistry of serial sections of human adenocarcinomatous colon also demonstrated upregulated expression of *Sox9* protein in an ectopic pattern that co-localized with increased nuclear localization of *β-catenin* (13). Upon first inspection, these data appear to be in direct conflict with the observation that low levels of *Sox9* might maintain a non-differentiated cellular phenotype, as overexpression of *Sox9* in non-transformed small intestine cell lines or colon cancer cell lines decrease proliferation and prevent differentiation into both secretory and absorptive lineages (39, 58). However, mRNA analysis of whole tissues and cell lines does not account for the level of *Sox9* expression in individual cells, which is of critical importance due to the dose-dependent behavior of *Sox* factors. A tumor expressing higher levels of *Sox9* mRNA might do so because there is an increase in the number of *Sox9* "low" stem cells and *Sox9* "sublow" progenitor cells over negative controls. In this case, there would be an increase in absolute mRNA levels of *Sox9* due to ectopic cellular expression that does not necessarily indicate an increase in *Sox9* expression in cells that express the gene under normal conditions.

In addition, downregulation of *Sox9* was shown to be sufficient to allow increased expression of *CEA* in colonic adenocarcinomas *in vitro*, while overexpression of *Sox9* was sufficient to prevent expression of *CEA* in the same cell lines (58). The observation that *Sox9* regulates *CEA* expression by colonic adenocarcinomas *in vitro* is

clinically relevant, as *CEA* is associated with anti-apoptotic effects in colon cancer that can lead to metastasis (91, 151).

Jay, et al. hypothesized that *Sox9* might have pro-apoptotic effects in colonic adenocarcinoma and demonstrated that induced expression of *Sox9* leads to an increase in apoptosis in colon cancer cell lines *in vitro* (58). These data supported the prior association of *CEA* with decreased apoptosis and the authors' demonstration of the ability of *Sox9* to indirectly inhibit the *CEA* promoter, leading to the conclusion that regulation of *Sox9* might be an attractive approach to mediating anti-apoptotic effects of *CEA* in aggressive colon cancers (58, 151). A mouse model exhibiting *Sox9*-knockout specific to the intestinal epithelium revealed a significant change in colonic epithelial morphology, with pronounced crypt hyperplasia and the formation of atypical, "villus-like" structures protruding into the lumen (10). Supporting previous observations that *Sox9* might play a pro-apoptotic role in colon cancer, *Sox9* knockout mice exhibited a high rate of spontaneous microadenoma formation (10). These emerging data regarding roles of *Sox9* in colon cancer suggest that the gene is not only important in the development and progression of tumors, but might also be exploited toward diagnostic and therapeutic ends.

Preliminary evidence suggests roles for Sox4 and Sox17 in the small intestine and colon

In addition to *Sox9*, limited yet intriguing evidence has emerged that suggests other *Sox* factors also have important regulatory roles in maintenance and differentiation in the intestinal epithelium. Interestingly, like *Sox9*, *Sox4* and *Sox17* have both been shown to differentially regulate *β-catenin* activity in colon cancer *in vitro*, leading to changes in cellular proliferation (64, 119). It was found that *Sox17* is downregulated in

spontaneous adenomas of APC-min mice, while *Sox4* is upregulated (119). Expression analysis was supported by data revealing that forced expression of *Sox17* in colon carcinoma cells resulted in a downregulation of *β-catenin/TCF* activity and a decrease in proliferation, while *Sox4* overexpression had the opposite effect (64, 119). These data show that individual *Sox* factors differentially regulate intestinal epithelial proliferation and suggest different ratios of *Sox* factors in the same cell may work in concert to control proliferation. Furthermore, the observation that both *Sox9* and *Sox17* negatively regulate *Wnt* activity is suggestive of overlapping roles for *Sox* factors in the intestinal epithelium.

PANCREAS

The study of pancreatic development and cellular homeostasis is arguably one of the most dedicated efforts to translate basic cell biology to clinically applicable therapies. The pancreas resides at an epicenter of questions concerning both individual quality of life and public health issues, as both type I and type II diabetes present with significant co morbidity at increasingly high rates in the general population, and are projected to affect 366 million individuals globally by 2030 (149). Of special interest is the idea that clinical replacement of absent or dysfunctional β-islet cells could greatly reduce the wide-sweeping impact of diabetes on the healthcare system as a whole.

Stem cells in the pancreas

The pancreas is a complex organ that fulfills both exocrine and endocrine functions. A majority of the pancreas is dedicated to exocrine function, producing enzymes essential for the digestion of macronutrients. The functional unit of the endocrine pancreas comprises specialized endocrine cells in the islets of Langerhans, which are interspersed throughout the exocrine pancreas. The islets consist of α-cells

(glucagon-secreting), β-cells (insulin-secreting), δ-cells (somatostatin-secreting), ε-cells (ghrelin-secreting), and PP-cells (pancreatic polypeptide-secreting). The regulation of pancreatic endocrine cell behavior is vital for homeostasis of the organism and an integral component in the pathogenesis of diabetes. As such, a great deal of attention has been devoted to understanding putative stem cell populations within the adult pancreas. To date, the precise location of pancreatic stem cells remains debated, but evidence suggests that ductal, acinar, and islet cells are capable of contributing to pancreatic regeneration (reviewed by (15). Strikingly, recent evidence has also shown that mature β-cells are capable of driving β-cell regeneration under both normal conditions and following injury (35, 90). The ability to modify proliferation strategies either intrinsically or in response to extrinsic signaling is suggestive of complex and context-specific transcriptional regulation of putative stem cell populations in the pancreas.

Sox factors have been an attractive target of research in pancreatic physiology and disease due to their ability to modulate the proliferative capacity of stem and progenitor cells. In terms of organs of the gastrointestinal tract, the breadth of *Sox* factors known to be expressed in the pancreas is rivaled only by that expressed in the intestine. Many *Sox* genes have been localized to the pancreas, including *Sox4*, *Sox6*, *Sox9*, and *Sox13* (see Table 1). (57, 62, 116, 150). *Sox4*, *Sox6*, *Sox9*, and *Sox13* are especially noteworthy in that they have been shown to have functional roles in pancreatic physiology. By contrast, the current knowledge of other *Sox* factors in the pancreas is restricted to expression analysis data, with no clear functional roles yet determined. This discussion of *Sox* factors in the pancreas will therefore focus on evidence for functional roles of *Sox* expression in the maintenance and differentiation of pancreatic cell populations.

Sox9 in pancreatic development

Sox expression in the developing pancreas is dynamic, with gene expression patterns and levels changing over the course of development. *Sox9*, the most studied Sox-factor in the pancreas, has been shown to exhibit dynamic expression throughout ontogeny. During murine development, *Sox9* undergoes changes in expression level that span broad expression at E9.5 to restricted, downregulated expression by late gestation (80). Spatial expression of *Sox9* has been shown to associate with the entire pancreatic epithelium at E10.5-E12.5, but at late gestation is restricted mostly to the islets of Langerhans (80, 116). Lineage-tracing studies have shown that cells expressing *Sox9* in the embryonic pancreas co-localize with multipotent progenitor cells that give rise to the pancreatic endocrine lineages, as well as to early pancreatic exocrine progenitors (115). These data show that *Sox9* marks a population of cells that give rise to all pancreatic cell types.

Sox4 shows similar changes in expression across pancreatic development as it is expressed at high levels early in development and undergoes downregulation between E12.5 and E15.5, but exhibits subsequent, slight upregulation in adult islets (80, 116). These studies support the notion that dynamic changes in *Sox* factor expression are a common theme in normal pancreatic development and suggest that *Sox* factors may regulate the timing of specific processes within the developmental timeframe.

Gene knockout experiments resulting in pancreas-specific ablation of *Sox9* in mice have demonstrated that *Sox9* is essential to pancreatic development, as knockout mutants died shortly after birth due to dehydration and elevated blood glucose levels, and displayed severely abnormal gross pancreatic morphology (116). In addition, pancreas-specific *Sox9* knockout mice exhibited heterogeneous phenotypes, with either a complete

absence of pancreatic endocrine cells or only sporadic representations of these cell types, suggesting that *Sox9* plays a critical role in the cellular development of the endocrine pancreas (116). Heterozygous mutants for pancreas-specific *Sox9* deficiency are viable and fertile and exhibit normal gross pancreatic morphology. Interestingly, upon histological examination, heterozygous mutants are shown to have significantly reduced numbers of pancreatic endocrine cells that are functionally normal in terms of mRNA expression of insulin processing enzymes and insulin secretion (115). These data suggest that while *Sox9* is important in terms of gross pancreatic development and early differentiation of pancreatic endocrine lineages, it is not essential for normal function of endocrine cells once they have undergone cell fate decisions. Furthermore, the authors note that varying degrees of deficit observed between haploinsufficiency and complete knockout are indicative of a dose-dependent role of *Sox9* in the endocrine pancreas (115). This notion is further supported by the observation that haploinsufficiency of *Sox9* in human patients suffering from campomelic dysplasia results in abnormal islet cell morphology (96).

Sox4 in pancreatic development

Similar to *Sox9*, expression patterns of *Sox4* are broad across pancreatic buds in early pancreatic development and restricted to the endocrine cells of the islets by adulthood (80, 150). Interestingly, observations regarding functional effects of *Sox4* have been very similar to those made for *Sox9*. Due to embryonic lethal cardiac defects in homozygous $Sox4^{null}$ mutants, all experiments examining pancreatic roles of *Sox4* past E11.5 were conducted *ex vivo* on explant cultures from mutant embryos (113, 150). There was no observed difference in growth rate between wild type, heterozygous, and homozygous, *Sox4* deficient explant cultures. Further examination revealed that

heterozygous mutant and wild-type cultures were able to generate islet cells *in vitro* that expressed insulin and glucagon, where homozygous mutants were not (113, 150). However, islet cells from *Sox4* homozygous knockout mice continued to exhibit normal markers of endocrine cell lineages, albeit at lower expression levels relative to heterozygous mutant and wild type explant cultures (113, 150). Decreased endocrine cell differentiation in $Sox4^{null}$ explant cultures suggests that *Sox4* plays a regulatory role in the expansion of the pancreatic endocrine compartment.

Possible functional roles of *Sox4* and *Sox9* in terms of regulating the cellular population of the endocrine pancreas are similar in that both seem to contribute to pancreatic cell number during development without affecting normal endocrine functions in cells that are able to properly differentiate. Wilson, et al conclude that *Sox4*, like *Sox9*, is a requisite component of normal endocrine expansion in the developing pancreas (113, 150). The similarities between observed functional roles of pancreatic *Sox4* and *Sox9* suggest that the two might play similar roles in the genesis of pancreatic endocrine cells. The regulation of pancreatic endocrine differentiation by *Sox4* and *Sox9* represents an area of research that warrants further investigation.

Sox6 is associated with hyperinsulinemia

In terms of functional maintenance of the adult pancreatic phenotype, *Sox6* is emerging as a potentially powerful mediator of gene regulation and cellular maintenance in pancreatic β-cells. Interest in pancreatic *Sox6* is supported by observations that persistent downregulation of *Sox6* is strongly associated with hyperinsulinemia in obese mice (56). Furthermore, downregulation of *Sox6* coincident with hyperinsulinemia was correlated with high-fat diet treatments in normal mice, as well as genetic defects leading to obesity (56). These data implicate *Sox6* in the dysregulation of insulin signaling

associated with type II diabetes mellitus. Whereas *Sox6* is shown to localize to normal adult β-cells and suppress genes implicated in β-cell proliferation, downregulation of *Sox6* has been shown to reduce this suppression and lead to an increase in glucose-stimulated insulin secretion (56). Furthermore, a reduction in levels of *Sox6* leads to proliferation and expansion of β-cell populations. Immunoprecipitation assays suggest that *Sox6* inhibits proliferation by forming complexes with β-catenin and serves as a cofactor for histone modifications that suppress downstream activities of β-catenin (57). Interestingly, a recent study utilizing bivariate genome-wide association study (GWAS) examination of polymorphisms present in obesity and osteoporosis in human males indicated that *SOX6* is strongly associated with both increased body mass index and decreased bone mineral density (81). These data would suggest that *Sox6* plays roles both in obesity-associated insulin regulation as well as regulation of β-cell proliferation in the adult pancreas.

Sox13 is a diabetes-associated autoantigen

Sox13 is expressed throughout pancreatic development and in adult islet cells, has been implicated in functional deficiencies in β-islet cells in type I diabetes mellitus (62, 80, 150). Surveys of type I and type II diabetes patients, as well as control subjects, demonstrated a higher occurrence of *Sox13* autoantigens in the diabetic population (62). Furthermore, *Sox13* autoantigens were significantly upregulated in type I patients compared to patients with type II diabetes (62). While these data demonstrate a strong correlation to autoimmune-associated type I diabetes, the authors are careful to point out that no definitive conclusions can be reached as to whether or not *Sox13* autoantigens are causative of, or consequential to type I diabetes.

Sox2 plays a role in the development of pancreatic cancer

In addition to roles in development and maintenance of pancreatic cellular phenotypes, *Sox* factors have been implicated in aberrant signaling processes leading to pancreatic cancer. As previously described in regards to gastric cancer, *Sox2* is also associated with abnormal cell types observed in pancreatic intraepithelial neoplasia (PanIN). Observations that *Sox2* is upregulated in PanIN correlate with the observed upregulation of extrapancreatic epithelial markers (100). *Sox2* has been shown to be essential for the maintenance of gastric epithelial phenotypes and abnormal expression of *Sox2* is associated with the development of gastric cancers through the failure to maintain this phenotype (92). Interestingly, levels of *Sox2* in PanIN are upregulated over normal pancreatic expression levels to levels normally seen in the gastric fundus and antrum (100). This, along with the identification of *pepsinogen C* and the gastric-specific mucin *Muc6* in PanIN, suggests that upregulation of *Sox2* in PanIN drives pancreatic cells to adopt an abnormal gastric epithelial phenotype (100). These studies continue to underscore the spatio-temporal and dose-dependent roles of *Sox* family transcription factors in maintenance and differentiation processes in the gastrointestinal tract.

HEPATOBILIARY SYSTEM

The hepatobiliary system, which consists of the liver, gallbladder, and bile ducts, plays a central role in proper function of the gastrointestinal tract. The liver serves a wide range of functions essential for both digestion and overall metabolism, acting as an exocrine organ as well as a filter for the systemic circulation. Homeostasis within hepatocyte populations, the functional parenchymal cells of the liver, is essential for proper hepatobiliary function and homeostasis of the entire organism. Pathologies

resulting in insufficient or aberrantly upregulated expansion of hepatocytes, such as cirrhosis or hepatocellular carcinoma, respectively, present clinical problems as well as significant losses in quality of life for patients.

Stem cells of the hepatobiliary system

The liver is compromised of individual functional units called hepatic lobules. Each lobule contains a portal triad consisting of a portal vein, hepatic artery, and bile duct, as well as hepatocytes, which are organized into rows called hepatic plates. Blood flows from the portal vein past the hepatocytes, which interact with the venous blood via a special fenestrated endothelium. After flowing past the hepatic plates, blood drains into the central vein, located opposite of the portal triad in the hepatic lobule (83). Normal hepatocyte turnover is slow, with hepatocyte lifespan having been calculated at approximately 200-300 days (18). Despite this low basal rate of turnover, the liver demonstrates remarkable regenerative characteristics following substantial loss of cells, as in the case of partial hepatectomy. In rats, the liver has been observed to regenerate and regain its original cell number in just 3-4 days following partial hepatectomy (17, 126). The rapid rate of regeneration lends itself to clinical benefit in liver transplantation, as it facilitates the use of living donors. Interestingly, radioactive thymidine-labeling studies have demonstrated that this tissue-regenerative process is driven by mature hepatocytes (17, 126). Stem cell populations in the liver do not appear to make significant contributions to regeneration, except in cases where hepatocytes are damaged, such as following the administration of toxic pharmacologic agents in animal models (36). When hepatocytes are significantly damaged, their regeneration appears to be dependent on a hepatic stem cell population. Progenitor cells of the liver, termed “oval cells”, are derived from a non-hepatocyte precursor and are classically viewed as bi-potential progenitors,

capable of giving rise to both hepatic and biliary epithelial cells (4, 144). The ability to undergo fate decisions between hepatic and biliary lineages is indicative of precise transcriptional regulation of genes controlling differentiation processes in hepatobiliary progenitor populations.

In addition, complex regenerative strategies of the liver suggest powerful molecular regulatory mechanisms that are capable of controlling regeneration differentially in homeostatic and pathologic states. Hepatocyte contribution to cellular regeneration following partial hepatectomy represents a scenario in which tight control of proliferative capacity would be essential in order to trigger mature cells to divide but also prevent aberrant growth once proper cell number was restored. The proven ability of *Sox* factors to regulate "stemness" in other tissues makes their potential roles in hepatobiliary stem and progenitor cell populations especially interesting. However, despite an increasingly detailed understanding of hepatic regeneration, the role of *Sox* factors in the hepatobiliary system remains poorly understood. Early research has revealed basic roles and produced expression data for a handful of *Sox* factors in the hepatobiliary system, including *Sox9* and *Sox17* (see Table 1). However, no single *Sox* factor has emerged as being central to hepatobiliary stem cell populations.

Sox9 and Sox17 contribute to biliary development

Recent research has begun to define emerging roles for *Sox9* and *Sox17* in the development of biliary ducts (7, 136). Biliary tubulogenesis proceeds via a two-step process, starting first with the differentiation of hepatoblasts into cholangiocytes followed by ductal morphogenesis, with cellular asymmetry (hepatoblasts v. cholangiocytes) between the parenchymal and portal regions of the developing duct maintained throughout tubulogenesis (7). Throughout the process of tubulogenesis, *Sox9* is

preferentially expressed on the portal/biliary side of the developing duct, with no apparent expression in the parenchymal/liver region from E10.5 onward. Postnatal expression of *Sox9* was found to be restricted to small ducts of the biliary system (7). Interestingly, liver-specific knockout of *Sox9* driven by *Alfp*-Cre did not result in aberrant biliary duct formation. Rather, *Sox9* knockout animals developed normal biliary ducts, but at a significantly slower rate than wild-type animals. Furthermore, *Sox9* was found to regulate the expression of other genes implicated in development, including genes in the *Notch* and *TGFβ* pathways (7). These data demonstrate that *Sox9* plays a critical role in the timing of biliary tubulogenesis, but also suggest that other compensatory mechanisms or signaling pathways exist that ultimately make *Sox9* non-essential for biliary duct development.

In contrast to *Sox9*, recent data suggests that *Sox17* is required for biliary development, specifically for the development of the gallbladder. Expressed throughout the foregut endoderm early in development, *Sox17* later becomes restricted to the gallbladder region of the foregut, between the hepatic and pancreatic buds (136). Interestingly, early expression of *Sox17* overlaps with markers for pancreatic fate, but with biliary induction later in development, *Sox17* is preferentially expressed in biliary primordium only (123). Tissue-specific knockout of *Sox17* results in improper positioning of the endoderm, consequently causing not only a lack of gallbladder development, but also resulting in ectopic development of pancreatic cells in the anatomical location of the gallbladder (123, 136). Conversely, continued expression of *Sox17* throughout development in cells also expressing early pancreatic markers resulted in the suppression of pancreatic development (123). These data demonstrate that *Sox17* is essential for the segregation and development of biliary primordium from early foregut

cells that give rise to both pancreatic and biliary cells. *Sox17* therefore appears to play an important role in fate specification during gallbladder and pancreatic development in the hepatobiliary system.

Preliminary findings point to wider roles for Sox factors in the hepatobiliary system

Aside from data generated by emerging biliary developmental research, *Sox* factors remain mostly uncharacterized in the hepatobiliary system, especially in terms of possible roles in hepatoblast maintenance and differentiation in the liver. *Sox2* has been shown to be upregulated in Hep-12 cells while *Sox7* and *Sox17* have been overexpressed in human embryonic stem cells to drive differentiation toward hepatic-like phenotypes in vitro, but functional roles of these *Sox* factors in normal hepatic development and hepatoblast maintenance remain undefined (114, 155). Additionally, *Sox13* has been identified as an autoantigen in primary biliary cirrhosis while *Sox2* is upregulated in tumors of the ampulla of vater (38, 108). Despite these findings, *Sox* factors in general are poorly described in pathogenesis in the hepatobiliary system. Data supporting powerful regulatory roles of *Sox* factors in stem cell populations of other gastrointestinal tissues suggests that the role of *Sox* factors in the liver is an area deserving of further investigation.

CONCLUSION

Diseases related to maintenance and differentiation of cells within the gastrointestinal tract, such as diabetes, colon cancer, and cirrhosis, affect a significant portion of the general population and often cause physiological consequences that negatively impact quality of life for affected individuals. An emerging understanding of tissue-specific stem cells holds significant promise toward the development of the next

generation of advanced clinical therapies, but the understanding of molecular regulation of these stem cells remains rudimentary. Initial studies have identified *Sox* factors as key players in the processes of stem cell maintenance and differentiation across nearly all organ systems in the gastrointestinal tract.

Sox factors possess attractive functional capabilities that allow them to serve as versatile and powerful regulators of transcriptional activity. To date, multidisciplinary research has yielded sufficient data to reach general conclusions concerning members of the *Sox* family of genes. In simplest terms, the *Sox* family is made up of genes that behave as classical transcription factors, but also have the ability to affect dramatic conformational changes through DNA bending. The ability to physically alter the shape of DNA has the potential to allow for the joining of distal enhancing elements with proximal transcriptional machinery, as well as allow for inhibitory effects.

Sox factors are attractive from a research standpoint in adult tissue due to their broad expression patterns, implication in maintenance of stem cell populations, and context-dependent functional roles. In large organ systems, such as the gastrointestinal tract, *Sox* factors have the potential to reveal unifying themes across organs that are functionally dissimilar, but derived from the same germ layer. Additionally, the feature of dose-dependent behavior might allow *Sox* factors to differentially modulate transcriptional responses to mitogenic signals, regulating proliferative potential in stem and progenitor cell populations.

From a translational standpoint, the ability to identify and manipulate cell populations via *Sox* function could lead to the directed control of stem cell populations *in vivo*. Additionally, understanding *Sox* function could allow for the *in vitro* expansion and autologous transplantation of gastrointestinal cells and tissues, a long-standing

therapeutic ideal of stem cell research. Dose-dependency and redundancy might allow *Sox* factors to be exploited in a manner that allows for the "fine-tuning" of healthy and diseased cells at the genetic level.

Despite recent advances in functional characterization of *Sox* factors, many important questions regarding the roles of specific SOX genes in the gastrointestinal tract remain. One significant area that merits increased investigation is the expression and role of *Sox* factors in hepatobiliary physiology. While members of the *Sox* family have been used to mark a hepatic lineage in embryonic stem cell differentiation studies and described in terms of hepatic angiogenesis, no data exists describing possible roles for *Sox* in terms of hepatocyte function and proliferation (85, 123). Additionally, the ability of *Sox* factors to dramatically bend DNA warrants further investigation toward the functional role of *Sox*-mediated structural and higher-order genomic changes in stem cell populations. The elucidation of transcriptional roles for *Sox* factors might be limited by conventional genetic assays. Standard techniques employed to detect direct downstream targets of transcription factors in proximal promoter regions could theoretically prove inefficient in detecting distal *Sox* target genes that are activated through DNA bending and therefore not amenable to this form of detection. The development of novel technical approaches to overcome this obstacle is vital to the further study of *Sox* factors and their role in stem cell populations. Evidence to date and the potential for translational application strongly supports further investigation into the role of *Sox* factors in tissue-specific stem cell populations of the gastrointestinal tract.

CHAPTER 3

SOX9 EXPRESSION MARKS A SUBSET OF CD24-EXPRESSING SMALL INTESTINE EPITHELIAL STEM CELLS THAT FORM ORGANOIDS IN VITRO

ADAM D GRACZ, SENDHILNATHAN RAMALINGHAM, SCOTT T MAGNESS

INTRODUCTION

The intestinal epithelium is one of the most proliferative tissues in the mammalian organism. The entire mono-layer of epithelium is replaced approximately every 7 days (12, 24), and this renewal process is driven by a pool of multipotent, self-renewing stem cells that are positioned between Paneth cells and also immediately above the Paneth cell compartment in the base of the crypt (8, 109). Typically, in the adult organism stem cells are thought to undergo an asymmetric division process that generates a larger pool of more rapidly dividing transit-amplifying populations (or progenitors) that are located higher up the crypt/villus axis towards the lumen (12, 24). Intrinsic and extrinsic signaling direct these progenitors to commit to one of four functionally distinct post-mitotic cell types: the enterocyte (absorptive) (22), goblet cell (mucous producing) (20), enteroendocrine cell (hormone producing) (23), and Paneth cell (antimicrobial peptide producing) (21). With the exception of the Paneth cell population, the other three cell types migrate up the crypt-villus axis and are sloughed into the lumen at the villus tip. Paneth cells are born around cell position +5 (as numbered from the base of the crypt towards the lumen) and migrate down to the crypt base where they secrete antimicrobial peptides into the crypt lumen (21). Normal gut homeostasis is critically reliant on the

proper control of stem cell proliferation, division, and subsequent generation of the transit-amplifying progenitor pool.

Significant progress towards identifying, isolating, and manipulating the IESC has been hindered by the lack of specific IESC biomarkers and suitable culturing techniques. Recently, a number of putative IESC biomarkers have been identified, but relatively few of these biomarkers have been tested to confirm that cells expressing these biomarkers pass the test of multipotency and self-renewal, the two functional characteristics that define a stem cell. Three studies have employed *Cre*-recombinase mediated genetic lineage tracing technology, currently the most rigorous method to assess multipotency *in vivo*, to demonstrate that the biomarkers, *Lgr5*, *Bmi1,* and *CD133/prominin1* are expressed in cell populations that demonstrate multipotency and self-renewal capacity in the small intestine (8, 109, 121). The expression patterns of *Lgr5*, *Bmi1*, and *CD133/prominin1* mark stem cell populations that are differentially positioned in distinct regions of the crypt. *Lgr5* expression is restricted primarily to the crypt-based columnar cells (CBC) that are intercalated between the Paneth cells (8). *Bmi1* also shows a highly restricted expression pattern that is primarily localized to the supra-Paneth cell region, the zone typically defined as cell position +4 (109). *CD133/prominin1* exhibits a broad expression pattern encompassing the CBC positions through at least cell position +10 (121). Although all three of these stem cell populations demonstrate multipotent capacity *in vivo*, it is unclear whether they have functionally equivalent roles in the homeostasis of the intestinal epithelial monolayer.

We have recently shown that distinct levels of *Sox9*-expression mark putative IESCs based on enriched levels of *Lgr5* mRNA in discrete cell populations (39). SOX9 is a transcription factor that not only marks stem/progenitor cells in various tissues, but it

also has an established role in maintaining the multipotent and proliferative capacity of stem/progenitor populations (88, 89). Using fluorescence activated cell sorting (FACS) of dissociated small intestine epithelium from a *Sox9*^{EGFP} reporter gene mouse model, we demonstrated that 'low' levels of *Sox9*^{EGFP} (termed *Sox9*^{EGFPlo}) mark cells that are enriched for *Lgr5* (39); moreover, we identified that 'high' levels of *Sox9*^{EGFP} (termed *Sox9*^{EGFPhi}) mark post-mitotic enteroendocrine cells (39). Expression of endogenous levels of *Sox9* mRNA and SOX9 protein were consistent with the *Sox9*^{EGFPlo} and *Sox9*^{EGFPhi} levels validating that the *Sox9*^{EGFP} reporter gene faithfully recapitulates expression patterns of endogenous *Sox9* (39). Although formal inducible *Cre*-mediated genetic lineage tracing studies using *Sox9* as a biomarker have not yet been conducted in an adult experimental model, *Cre*-mediated lineage tracing during embryonic gut development suggests that *Sox9*-expression marks multipotent IESCs at E17 (1). This embryonic lineage tracing study provides compelling evidence that *Sox9*-expression marks a population of IESCs in the embryo that may be conserved in the adult.

Due to the CBC location and enriched *Lgr5* expression in the *Sox9*^{EGFPlo} population, we hypothesized that *Sox9*^{EGFPlo} expression marks functional IESCs. To test this hypothesis we utilized a recently described novel *in vitro* assay that provides a culturing environment, which supports the generation of crypt-villus like structures from a single *Lgr5*-expressing intestinal epithelial stem cell (112). In this study we also characterize genetic signatures of small intestine epithelial stem/progenitor populations exhibiting different *Sox9*-expression levels, and explore the multipotent and self-renewal capacity of these different populations of *Sox9*-expressing crypt cells. In an attempt to identify a cell surface marker that can be used to enrich for multipotent IESCs without the requirement of a specialized fluorescent reporter gene animal model, we investigate the co-expression

pattern of *Sox9*^{EGFPlo} and the Cluster-of-Differentiation marker, *CD24*, a cell surface marker that has been reported to be putative stem cell biomarker and is easily detected using widely available FACS antibodies (104, 145).

MATERIALS AND METHODS

Mice/Genotyping

The *Sox9*^{EGFP} mouse line was originally generated as part of the GENSAT Brain Atlas Project (43) and contains genomic integration of a modified BAC (RP32-140D18) with ~75.5 kb upstream and ~151kb downstream sequence to *Sox9*. Frozen *Sox9*^{EGFP} mouse embryos were obtained from the Mutant Mouse Regional Resource Center (University of California-Davis) and reconstituted by transfer into foster mice. All mice are on the outbred CD-1 strain and were maintained as heterozygotes on the CD-1 genetic background. Mice breed normally and live to adulthood with no overt phenotypes due to the transgene. At ~10 days post-natal, tail snips were viewed under an epi-fluorescent microscope fitted with filters for EGFP visualization. A high level of EGFP fluorescence compared to transgene negative control mice was scored as positive for the *Sox9*^{EGFP} transgene.

Tissue dissociation/FACS

To isolate intestinal crypt cells for Fluorescence Activated Cell Sorting (FACS), small intestine epithelium was dissociated into single cells essentially as previously described (33) with the following modifications. For FACS experiments mouse intestines were flushed with cold phosphate-buffered saline (PBS), cut open lengthwise in approximately 10 cm long pieces, and immersed in PBS/30 mM ethylenediaminetetraacetic acid (EDTA)/1.5 mM dithiothreitol (DTT) over ice for 20 minutes. The solution was disposed of and the tissue was shaken vigorously in fresh

PBS/30 mM EDTA for approximately 30 seconds before being incubated at 37°C for 10 minutes. Intact tissue was discarded and dissociated crypts and villi were pelleted at 2500 rpm for 5 minutes. The cells were washed twice with cold PBS, resuspended in Hank's buffered saline solution (HBSS)/0.3U/mL dispase at 37°C and shaken approximately every 2 minutes for 10 minutes. Then, fetal bovine serum (FBS, 10%) and 100 μg DNaseI was added before the cells were passed through a 100 μm filter. Cells were pelleted at 2500 rpm for 5 minutes and resuspended in 4 mL HBSS with 10% FBS, then passed through a 70 μm filter and combined with an additional 100 μg DNaseI. Equivalent numbers of cells from three animals were combined and passed through a 30 μm filter immediately prior to FACS.

Sox9^{EGFP} and CD24–expressing cells were isolated using a MoFlo XDP FACS machine (Dako/Cytomation, Carpinteria, CA). Cells were collected in ice cold Advanced DMEM/F12 supplemented with 10μM Y27632 (Sigma, Saint Louis, MO) and kept on ice throughout the sort. Forward-Side Scatter gating was used to exclude 99.8% of all dead cells and lymphocytes. Both Forward Scatter and Side Scatter height-width plots were used for doublet discrimination to maximize efficient single cell sorting. Sort gates for *Sox9*^{EGFP} expression were defined by previous studies (39) with the inclusion of more precise parameters to include the *Sox9*^{EGFPsubLo} population. Cells were sorted into medium described above.

For CD24 or CD133 staining $1x10^7$ cells were isolated as described above and were stained with 5μL rat anti-mouse *CD24* preconjugated with Pacific Blue (BioLegend, San Diego, CA) or anti-mouse CD133 preconjugated with Allophycocyanin (APC) (BioLegend) in 2mL Advanced DMEM/F12 (Invitrogen, Carlsbad, CA) for 1 hour on ice. The cells were washed twice in Advanced DMEM/F12 prior to FACS. Gate decisions for

CD24 or CD133 positive cells were made based on comparison to Pacific Blue or APC conjugated isotype control stained cells.

Tissue Culture

Tissue culture methods were carried out essentially as previously described (112) with the following modifications. FACSorted cells were transferred to 2.0 mL conical tubes and pelleted at 4500 rpm for 5 minutes at 4°C before being resuspended at an approximate density of 2000 cells/50μL/well (24-well plate) in Matrigel (BD Biosciences, San Jose, CA) supplemented with 1μM 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). All mass per volume growth factor concentrations were calculated in respect to the 500μL final volume of media per well. To facilitate ease of observation, 50μL droplets of Matrigel/cell suspension were added per well to 24-well plates. After total polymerization, each formed droplet was overlaid with 500μL Advanced DMEM/F12 containing N2 supplement (Invitrogen), B27 supplement minus vitamin A (Invitrogen), 10mM HEPES (Invitrogen), and 10μM Y27632 (Anoikis inhibitor). Growth factors were added every other day at the same initial concentrations, with the exception of R-SPONDIN1, the dosage of which was reduced to 500ng/mL for all time points following the initial plating. Media was replaced every four days. Y27632 was withdrawn at culture day four with the first media change and not included in any subsequent culture maintenance.

For passaging experiments, organoids were isolated from Matrigel with a pipette and transferred to Advanced DMEM/F12 with 10μM Y27632 in 2.0mL conical tubes. Single cells were obtained by mechanical trituration of isolated organoids followed by incubation in 0.3U/mL Dispase (BD Biosciences) for 20 minutes at 37°C combined with

trituration every 5 minutes before media was collected and cells were plated again in Matrigel as described above.

cDNA Preparation/Real-time PCR Analysis

cDNA from approximately 0.75-1.0 x 10^5 cells from each FACSorted population (*Sox9*^{EGFPneg}, *Sox9*^{EGFPsubLo}, *Sox9*^{EGFPlo}, and *Sox9*^{EGFPhi}) was made using RNAqueous Micro Kit (Ambion, Austin, TX) according to the manufacturer's protocols. Time between the death of the mouse to RNA extraction was kept to 3.5 to 4 hours to ensure the highest quality of RNA. Real-time PCR was conducted for each sample in triplicate on approximately 1/20,000 of the total amount of cDNA generated. Taqman probes (*18S*, HS99999901; *Ascl2* Mm01268891_g1; *Atoh1,* Mm00476035_s1; *CD133* Mm00477115_m1; *CD24* Mm00782538_sH; *ChgA* Mm00514341_m1; *cMyc* Mm00487803_m1; *CyclinD1* Mm03053889_s1; *S100a4 (FSP), Mm00803371_m1; Hes1*, Mm00468601_m1; *Lactase* Mm01285112_m1; *Lgr5*, Mm00438890_m1; *Notch1*, Mm00435245_m1; *Olfm4* Mm01320260_m1; *Smooth Muscle Actin (SMA),* Mm01546133_m1; *Sox9*, Mm00448840_m1; *Substance P,* Mm00436880_m1) for each gene were obtained from Applied Biosystems (Pleasanton, CA) and used in reactions according to the manufacturer's protocol. 18S ribosomal RNA was amplified and used as the internal control gene for sample comparison. Delta CT values were calculated to obtain fold-changes for sample comparison (95).

Immunostaining/Microscopy

For tissue preparation, small intestines were dissected from adult *Sox9*^{EGFP} mice (>8 weeks of age) and luminal contents were flushed-out with PBS followed immediately by a single flush with freshly made 4% paraformaldehyde (PFA). The intestine was opened along the duodenal-ileal axis, and fixed for an additional 14-18 hours at 4°C. The

tissues were then prepared for cryosectioning by immersion in 30% sucrose solution for at least 24 hours at 4°C. Tissues were then embedded in Optimal Cutting Temperature (OCT) medium and frozen on dry ice. The OCT blocks were stored at -80°C until cryosectioning. Thin sections (8-10μm) were cut on a cryostat and placed on positively charged microscope slides for staining and microscopy. This tissue preparation technique is critical for preserving the EGFP fluorescence.

Organoids were prepared for immunostaining by aspirating culture media from Matrigel cultures and fixing entire contents of each well with 4% PFA for 14-18 hours at 4°C. Following fix, PFA was removed and replaced with 30% sucrose solution for at least 24 hours at 4°C. Organoids were then removed from Matrigel with a micropipette, embedded in OCT medium, and frozen on dry ice. The OCT blocks were stored at -80°C until cryosectioning. Thin serial sections (8-10μm) were cut on a cryostat and placed on positively charged microscope slides for staining and microscopy.

For immunostaining, the sections were washed twice in PBS to remove OCT, followed by incubation in blocking medium (5% normal goat or donkey serum (NGS or NDS)), in PBS-0.3% Triton-X100) for at least 30 minutes at room temperature (21-25°C). Primary antibodies were applied to the tissue sections in antibody staining solution (1% NGS, in PBS-0.3% Triton-X100). Dilutions were as follows: αCD326/EpCAM (rat, 1:1000, BioLegend, 118211), aLysozyme (rabbit, 1:1000, Diagnostics Biosystems, Pleasanton, CA, RP 028), aMucin2 (rabbit; 1:100, Santa Cruz Biotechnology, Santa Cruz, CA, sc-15334), aSOX9 (rabbit, 1:1000, Chemicon, Temecula, CA, AB5535), aSubstance-P (rat; 1:500, Chemicon, MAB356), aSucrase Isomaltase (goat; 1:100, Santa Cruz Biotechnology, sc-27603). All secondary antibodies (aRabbit-Cy3, Sigma, Saint Louis, MO, C2306; aRabbit-Alexafluor 488, Molecular

Probes, Eugene, OR, Z-25302; aRat-Alexafluor 555, Molecular Probes, Z-25305; aGoat-Cy3) were used at a 1:500 dilution in staining buffer. Nuclei were stained with Draq5 (1:10,000, Biostatus, Ltd, San Diego, CA, BOS-889-001). Background staining was negligible as determined by non-specific IgG staining.

Day to day development of organoids was tracked using light microscopy. Defects in culture plate plastic were utilized as "landmarks" to facilitate consistent observation of developing organoids. Epifluorescence images were captured on an Olympus IX70 fitted with an Olympus digital camera. Objective lenses used were 20X and 40X with numerical apertures of 0.55 and 1.40 respectively. All confocal images represent 1.0 μm optical sections unless otherwise noted. Objective lenses for the confocal images were 40X – 63X with a numerical aperture = 0.1.

RESULTS

A cell population enriched for IESC biomarkers can be distinguished from progenitor cells based on Sox9^{EGFP} levels

Sox9 is expressed at different levels in cells localized to the bottom region of the crypts (39, 140). Our previous study demonstrated that high levels of *Sox9* marked mature enteroendocrine cells, and that lower levels of *Sox9* marked a cell population that was enriched for the experimentally validated biomarker, *Lgr5*. Our initial report on *Sox9* expression in the small intestine crypts did not describe a third population that can be visualized by immunostaining for endogenous SOX9 and also by *Sox9^{EGFP}* transgene expression (Figure 1A). Distinct *Sox9^{EGFPsubLo}* cells are observed by immunostaining and transgene expression in all crypts of the small intestine (data not shown). We have defined these cells as 'sub-low' (subLo) expressing cells based on low but detectable levels of SOX9 and *Sox9^{EGFP}*, and show that these cells localize to the transit-amplifying

progenitor cell population (Figure 1A and S1). Using the mouse model in which *Sox9* regulatory regions control expression levels of the EGFP reporter gene, we sought to define parameters that would allow specific separation and isolation of IESCs from progenitor populations. We reasoned that by further dividing low-expressing $Sox9^{EGFP}$ cells into two more precise categories: a) $Sox9^{EGFPlo}$ levels, and b) the lowest detectable levels of $Sox9^{EGFP}$, termed $Sox9^{EGFPsubLo}$, we would potentially enrich and separate $Sox9^{EGFPlo}$-expressing IESCs from $Sox9^{EGFPsubLo}$ transit amplifying cells (Figure 1B). FACS cells based on $Sox9^{EGFPneg}$, $Sox9^{EGFPsubLo}$, $Sox9^{EGFPlo}$, and $Sox9^{EGFPhi}$ sort parameters demonstrated efficient separation of the 4 populations of $Sox9^{EGFP}$-expressing cells (Figure 1C).

To determine the genetic profiles of each of these populations, we conducted real-time semi-quantitative RT-PCR (sqRT-PCR) for 'stemness' and lineage specific genes. *Sox9* mRNA levels were consistent with $Sox9^{EGFP}$ transgene levels, thus validating the $Sox9^{EGFP}$-based FACSort (Figure 2A). $Sox9^{EGFPlo}$ cells show enriched expression for the validated IESC biomarker, *Lgr5*, as well as the two putative IESC biomarkers, *Olfm4* and *Ascl2*, which have been shown to co-localize with the *Lgr5* CBC population (139, 140) (Figure 2A). Two hallmarks of stem/progenitor cell maintenance are active Wnt/ß-catenin (71) and Notch-pathway signaling (103, 141). To determine whether these two pathways were active in $Sox9^{EGFPlo}$ cells we assessed the expression levels of *Wnt*-target genes, *c-Myc* and *CyclinD1*, and also *Notch1* and its downstream target, *Hes1* (Figure 2B). Our result show increased expression of *c-Myc/CyclinD1* and *Notch1/Hes1* in the $Sox9^{EGFPlo}$ and $Sox9^{EGFPsubLo}$ populations demonstrating intact signaling of these two pathways in both populations. Interestingly, $Sox9^{EGFPsubLo}$ cells were distinguishable from $Sox9^{EGFPlo}$ cells by marked up-regulation of *Atoh1* (*Math1*) (Figure 2C). *Atoh1*-

expression is an indicator of early stem cell differentiation into secretory lineages (156), consistent with a 'progenitor cell' nature of the *Sox9*^{EGFPsubLo} population. Analysis of the lineage specific genes, *Chromogranin A* and *Substance P*, show that the *Sox9*^{EGFPhi} sort parameters enrich for enteroendocrine cells. Increased *Lactase* expression in the *Sox9*^{EGFPneg} population indicates this population is enriched for differentiated enterocytes (Figure 2C). Together these results demonstrate that FACS based on *Sox9*^{EGFP}-expression levels allows for the efficient isolation and enrichment of distinct cell populations of the small intestine epithelium, specifically, *Sox9*^{EGFPhi} enteroendocrine cells, *Sox9*^{EGFPlo} IESCs, *Sox9*^{EGFPsubLo} progenitors/transit amplifying cells, and *Sox9*^{EGFPneg} differentiated enterocytes.

Single Sox9^{EGFPlo} intestinal epithelial stem cells generate organoids and differentiate into Paneth cells, enteroendocrine cells, goblet cells, and enterocytes

Sox9^{EGFPlo} cells show enriched expression for established markers of multipotent IESCs, but we wished to establish if these cells possess functional multipotent characteristics of IESCs. To assess whether *Sox9*^{EGFPlo} cells have the capacity to give rise to crypt/villus-like structures containing all four post-mitotic epithelial cell populations of the small intestine, we dissociated small intestine epithelium from *Sox9*^{EGFP} mice, isolated *Sox9*^{EGFPlo} cells by FACS, and plated cells into a 3-dimensional (3-D) culture system that has been recently described (112). A single cell was identified and followed using bright field and fluorescence microscopy from 24-hours post-plating to 19-days post plating to assess growth rate and morphological changes (Figure 3 and S5). Twenty-four hours post-plating, the single cells undergo cell division generating micro-colonies, which are composed of cells that maintain *Sox9*^{EGFP} expression (Figure 3B). By 48-hours post-plating the micro-colony develops a pseudo-lumen that is fully developed by 6-days

post-plating. The pseudo-lumen contains dead/apoptotic cells, which was indicated by punctate, fractured nuclear staining (data not shown), and high levels of autofluorescence. Clear segregation of cell types expressing distinct levels of *Sox9*^{EGFP} was observed by day 9 and was coincident with the initial formation of crypt buds, which were fully developed by day 12-16 (Figure 3C). Two gross morphologies of 'organoids' (112) were observed, one defined by an 'open' configuration where the dead/apoptotic cells were expelled into the Matrigel, and a second defined by a 'closed' confirmation where the dead/apoptotic cells were expelled into a central luminal cavity (Figure 3C). Organoids were allowed to grow in culture until they either became disorganized or grew large enough to come into contact with the bottom of the plate, which usually occurred between 16 and 20 days in culture. At this point, the organoids were dissociated by mechanical trituration and re-plated into the 3-D culture system. The re-plated cells were able to give rise to multiple organoids for multiple passages (at least 6 to date) indicating the self-renewal capacity of the IESCs within the original organoid. Single *Sox9*^{EGFPlo} cells were able to give rise to organoids with a well-defined pseudo-lumen (day 2-6) at an average incidence of 5% (5 organoids/100 FACSorted *Sox9*^{EGFPlo} cells plated), however, fully-formed organoids with well-defined crypt units that were present at day 12-16 were more rare and were present at an average incidence of 1%. It is important to note that upon culturing *Sox9*^{EGFPneg}, *Sox9*^{EGFPsubLo}, and *Sox9*^{EGFPhi} cells mature organoids never formed, further validating that *Sox9*^{EGFPlo} sort parameters were effectively isolating the multipotent, self-renewing IESC population.

To determine whether *Sox9*^{EGFPlo} cells in the organoid where able to give rise to the four post-mitotic differentiated cells types of the small intestine epithelium, we sectioned a single organoid and used immunostaining to identify the cell lineages

represented. The data indicate cells in the organoid are epithelial in nature as demonstrated by positive immunostaining with Epithelial Cell Adhesion Molecule (EpCAM; Figure 4A). The three secretory lineages were represented and identified by immunostaining for Lysozyme (Paneth cell; Figure 4B); Substance P (Enteroendocrine; Figure 4C); and Mucin2 (Goblet; Figure 4D). An interesting observation was that high levels of Mucin2 were present in the pseudo-lumen, which is physiologically consistent with normal gut function, and furthermore indicates cellular polarity. Enterocytes were identified by immunostaining for the brush-border enzyme, Sucrase Isomaltase (SIM) (Figure 4E). These data indicate that single *Sox9*^{EGFPlo} IESCs have the capacity to function as multipotent IESCs in culture.

One pressing question is whether contaminating mesenchymal cells, which may be attached to a *Sox9*^{EGFPlo} IESCs and therefore undetectable using FACS doublet-discrimination parameters, are providing instructive signals and mitogens to the organoids. To address this question, we visually assessed post-sort cells using fluorescence microscopy and never identified a cell aggregate containing a non-*Sox9*^{EGFP}-expressing cell. A more objective approach to answer this question was to section an entire organoid and immunostain for biomarkers of fibroblasts (Fibroblast Specific Protein-S100a4) and myofibroblasts (smooth muscle actin – SMA), the major cell types that have been proposed to support the stem cell *niche* and crypt formation *in vivo* (68). Although tissue sections of intact small intestine demonstrated robust immunostaining staining for fibroblast/myofibroblast markers, no staining was ever observed in 12-16 day old organoids indicating cells marked by SMA or S100a4 do not appear to be supporting *in vitro* growth of organoids (data not shown).

CD24 can be used to isolate and enrich for IESCs that form organoids in culture

To date, no cell surface marker has been identified that can be used to FACS-isolate normal IESCs. Although investigators have been searching for a single biomarker that exclusively marks stem cell populations, no one single gene has been demonstrated to possess this characteristic. Prevailing evidence suggests that pure stem cell populations may be isolated from a larger pool of heterogeneous cells by using a combination of cell surface biomarkers that produce a genetic signature for the stem cell. For example, Cluster-of-Differentiation (CD) markers are a class of cell surface proteins that have been used to successfully FACS-isolate multipotent stem cells from other epithelial tissues such as the brain, pancreas and mammary glands (3, 101, 104, 145). One common feature of all these epithelial stem cell populations is that CD24 is expressed on the cell surface and can be used in combination with FACS to enrich for stem cells from the heterogeneous cell isolates. CD24 is a short signal-transducing cell surface protein that was first characterized in B-lymphocytes (54, 65, 66) and has been since shown to be a ligand for P-SELECTIN, an adhesion molecule found on activated platelets and endothelial cells (107). The role of CD24 in the small intestine epithelium is currently not understood.

To explore the possibility of using CD24 as potential enrichment factor for IESCs, we conducted gene expression analysis for CD24 using real-time semi-quantitative RT-PCR (sqRT-PCR) on $Sox9^{EGFPlo}$ and $Sox9^{EGFPsubLo}$ FACSorted cells and compared their CD24 expression with CD24 expression in $Sox9^{EGFPneg}$ cells (Figure 5A). The data show there is a 5- to 7-fold increase in *CD24* mRNA in $Sox9^{EGFPsubLo}$ and $Sox9^{EGFPlo}$ cells compared to $Sox9^{EGFPneg}$ cells (Figure 5A). Likewise, flowcytometric analysis demonstrates CD24 protein is expressed in >99% of $Sox9^{EGFP}$ positive cells

(Figure 5B). These data support the concept that *CD24* is a potential candidate biomarker for isolating and enriching IESCs from non-transgenic tissue sources.

CD24 is present on the surface of some non-IESC types; therefore, FACS gates must be clearly defined to specifically identify where the IESC population is located on a CD24/Forward-Scatter FACS histogram. Since organoids could only be generated from IESCs isolated using $Sox9^{EGFPlo}$ parameters, we reasoned that $Sox9^{EGFPlo}$ status could be used as a color-gating tool to identify where IESCs were located on a FACS histogram with only CD24 and forward-scatter as variables. Color gating is an algorithmic feature of flowcytometry software that allows a defined parameter (like EGFP expression status) to be mapped back onto cells on a FACS histogram that does not contain an EGFP variable. In other words, $Sox9^{EGFPlo}$ cells can be recognized on a CD24 vs. forward-scatter histogram to reveal where the $Sox9^{EGFPlo}$ IESC population falls within CD24-expressing cells. The significance of using this strategy is that defined sort parameters (or 'sort-gates') can be specifically drawn to exclude the majority of CD24-expressing cells that do no contain IESCs. The IESC-specific CD24 gates can then be used as a standard to isolate and enrich for IESCs from any mouse intestinal epithelial tissue source, not just from specialized fluorescent reporter gene mice.

$Sox9^{EGFPlo}$ intensities (Figure 6A-red events) that produced enrichment of IESCs and generation of organoids were used to define sort-gates on a bivariate histogram of CD24-stained intestinal epithelial cells from a $Sox9^{EGFP}$ mouse (Figure 6C-F, red-events). A bivariate *EGFP*/CD24 histogram indicated there were four discernable populations, which were used to define sort gates (CD24-high, medium, low, and negative) (Figure 6B). Color-gating only the $Sox9^{EGFPlo}$ cells onto the CD24/FSC histogram shows that a majority (60%) of all $Sox9^{EGFPlo}$ cells fall into the $CD24^{lo}$ gate (Figure 6D, F; red events)

representing 19% of all cells within the CD24^{lo} sort parameter (Figure 6F & 7A). CD24^{med} and CD24^{hi} gates contain significantly less *Sox9*^{EGFPlo} cells, 9% and 4%, respectively. A significant number of *Sox9*^{EGFPlo} cells (27%) fell in the CD24^{neg} sort parameter, however, this represented only 0.5% of all cells within that gate indicating a de-enrichment of IESCs in the CD24^{neg} population (Figure 6D, F & 7A).

As proof-of-concept, the same color-gating strategy was used to determine whether CD133, a validated IESC marker by lineage tracing, would also be a candidate cell surface marker for the isolation of *Sox9*^{EGFPlo} IESCs. CD133 has a broad expression pattern that encompasses the CBC cells up the crypt axis to the transit-amplifying cell zone (121). sqRT-PCR and flowcytometric analyses indicate that CD133 mRNA and protein are expressed in the *Sox9*^{EGFPsubLo} and *Sox9*^{EGFPlo} populations (Figure S2). Color-gating demonstrates that the *Sox9*^{EGFPlo} population (94.8%) expresses CD133 on the surface, however, the CD133-APC fluorescence shift is too low to enable significant separation of the *Sox9*^{EGFPlo} IESCs from CD133–negative cells (Figure S3). If the sort gates indicated on the CD133/FSC histogram were used to isolate *Sox9*^{EGFPlo} IESCs, only 10.7% of the cells would be *Sox9*^{EGFPlo} cells representing approximately half the enrichment factor compared to using CD24 as a FACS marker.

To identify which CD24 gate contained cells with functional IESC characteristics, we introduced the four cell populations (CD24 negative, high, medium and low) into organoid culture conditions. Organoids were generated only from CD24^{lo} and CD24^{med} populations at an incidence of 1:200 (CD24^{lo}) and 1:1000 (CD24^{med}) (Figure 7A, B). There was no overt difference in the morphology or growth between organoids derived from cells isolated from the CD24^{lo} and CD24^{med} gates. No organoids formed from CD24^{neg} or CD24^{hi} populations indicating the cells in these two populations do not have

stem cell capacity in culture conditions. Organoids derived from $CD24^{lo}$ and $CD24^{med}$ cells demonstrated multipotency (Figure S4) and have been able to be expanded in culture for at least 4 passages to date.

DISCUSSION

In this study we report that $Sox9^{EGFPlo}$-expression levels can be used to isolate intestinal epithelial stem cells (IESCs) that are functionally defined by their ability to form organoids *in vitro*, self-renew, and contain all four post-mitotic cell types of the small intestine epithelium. Additionally, we show that distinct levels of *CD24* expression enrich for IESCs, and can be used to efficiently FACS and isolate IESCs from non-transgenic intestinal tissue, providing to our knowledge the first reported cell surface marker that can be used to FACS isolate and enrich for IESCs with functional characteristics.

Although we have previously reported two distinct levels of *Sox9* expression in the small intestine epithelium (39), in this study we describe a third level of *Sox9*-expression termed $Sox9^{EGFPsubLo}$. Co-localization of $Sox9^{EGFPsubLo}$ cells with the transit-amplifying zone of the crypt, FACS enrichment for genetic markers of dividing progenitor cells, and the lack of multipotent and self-renewal capacity of $Sox9^{EGFPsubLo}$ cells suggest that sub-low levels of *Sox9* mark transit amplifying progenitors. Using $Sox9^{EGFPsubLo}$ expression status as a cell isolation parameter will facilitate future studies on the small intestine epithelial progenitor population.

An emerging concept in the literature is that Sox-transcription factors function in a dose-dependent manner to affect competence, potency and proliferation of stem and progenitor cell populations (102, 115, 132). Results from this present study and our previous work indicate that high levels of *Sox9* are associated with non-dividing cells,

and a decreasing gradient of *Sox9* is associated with stem/progenitor populations with increased proliferative capacities (39). The control of proliferation by *Sox9* is in part exerted through modulating *Wnt/ß-catenin* signaling (2, 10, 39). *Sox9* has been reported to be a downstream target of *Wnt*-signaling and appears to function by suppressing *Wnt*-signaling in a negative feedback loop mechanism (2, 10). High levels of transiently-expressed SOX9 in a non-transformed intestinal epithelial cell line halt proliferation indicating that SOX9 influences expression of proliferation genes (39), however, the exact mechanism by which SOX9 exerts control of proliferation is not understood and warrants further investigation into downstream effector genes. The ability to specifically sort the more rapidly dividing transit amplifying progenitor cells from the slower dividing IESCs of the small intestine epithelium will facilitate these investigations.

The seminal study describing the culture conditions for supporting the development of IESCs *in vitro* reports that single *Lgr5*-expressing cells are capable of building crypt-villus structures without a mesenchymal *niche* (112). A fibroblastic cell component has traditionally been thought to be essential for crypt-villus development (67, 69, 99, 124); therefore it was interesting that stereotypical crypt-villus units could form in the absence of a supporting cellular *niche.* The 3-dimensional culture conditions that facilitate crypt-villus development in culture do however provide many critical components similar to those a mesenchymal cell might contribute to an IESC *niche*, including laminin enriched extra-cellular matrix (110), growth factors, and *Wnt*-agonists (112). Under culture conditions that mimic an *in vivo* IESC *niche*, $Sox9^{EGFPlo}$- and *CD24*-expressing IESCs do not appear to require instruction from a mesenchymal cell component in order to expand and build crypt-villus units in culture. Immunostaining for Smooth Muscle Actin (SMA) or Fibroblast Specific Protein (FSP) failed to demonstrate

positive staining in serial sections through *Sox9*^{EGFPlo} IESC derived organoids (data not shown). To rule out the possibility that identification of a small population of fibroblasts/myofibroblasts was beyond the threshold of detection of the immunostaining technology, we conducted sqRT-PCR on RNA isolated from both isolated crypts and post-sorted *Sox9*^{EGFPlo} cells. sqRT-PCR analysis also failed to detect SMA or FSP from either sample after 40 cycles of PCR indicating there is no detectable contamination of a fibroblastic cell in the FACSsorted cells (data not shown). These data provide strong evidence that generation of crypt-villus units derived from *Sox9*^{EGFPlo} cells do not require a fibroblastic cell component. This is consistent with the proposal that in this 3-D culture system IESCs are able to set-up a crypt-villus axis by differential intrinsic responsiveness of intestinal epithelial cells to *Wnt*-signals rather than from extrinsic stimuli (112).

To date, FACS single IESCs that express experimentally validated markers of IESCs has required a specialized transgenic reporter-gene mouse expressing a fluorescent protein (8, 39). A previously reported method for isolation of a putative IESC population from non-transgenic mice was via side population (SP) sorting (32). Although the SP faction was shown to be from the crypt base, and to be enriched for putative IESC markers (47), SP cells have not been demonstrated to be capable of either replication or differentiation. We report here that *CD24* is a cell surface biomarker that can be used to FACS enrich for cells with functional IESC characteristics (multipotency and self-renewal) and possess the ability to form crypt/villus-like units *in vitro* from a non-transgenic mouse. Low *CD24* expression has been used as a marker to FACS isolate stem cells from the pancreas, brain, and mammary tissue (3, 101, 104, 107, 145). The expression pattern of *CD24* in the normal small intestine epithelium is unknown, however, there are reports that *CD24* is associated with epithelial cells in colonic

adenocarcinomas (79, 105). When using *CD24* as a FACS biomarker for IESCs, organoids only developed from cells isolated from $CD24^{lo}$ and $CD24^{med}$ gates with an incidence of 0.5% and 0.1% respectively. This represents 2- to 10-fold reduction in organoid formation when compared to using $Sox9^{EGFPlo}$ expression alone as the sort parameter.

Recent lineage tracing studies in mice suggest that *CD133*-expression marks a sub-set of crypt cells that exhibit IESC characteristics (121). The use of CD133 as a cell surface marker for FACS enrichment of these IESCs was not tested. Our data suggest that CD133 is not an efficient cell surface marker for enrichment of murine IESCs with antibody reagents currently available (Figure S3). Co-staining with CD24 and CD133 with the goal of identifying a sub-population of putative IESCs failed to produce a unique fluorescent signature that could be used to further FACS enrich $Sox9^{EGFPlo}$ IESCs (data not shown). Putative IESCs marked by low levels of $Sox9^{EGFP}$ did exhibit expression of *CD133* mRNA and protein (Figure S2), however, flowcytometry indicated that CD133 expression levels appeared to be below a threshold that would allow separation and enrichment by FACS (Figure S3). A low avidity of the antibody for CD133 is an alternative explanation that might explain the low fluorescent signature of CD133 in the flowcytometry histogram. Additional antibody reagents must be developed to test this hypothesis.

Although the most rigorously validated biomarkers *Lgr5*, *Bmi1*, *CD133* and *Sox9* appear to mark IESC populations, none of these markers have been shown to exclusively mark IESCs. For instance, like $Sox9^{EGFP}$, $Lgr5^{EGFP}$ is expressed at high and low levels in the crypt-base. It was demonstrated that low-expressing $Lgr5^{EGFP}$ cells rarely gave rise to organoids (112) suggesting that *Lgr5* also marks a non-IESC population. Furthermore,

only 6% of *Lgr5*^{EGFPhi} cells demonstrated self-renewal and multipotent capacity (112). Although the low number *Lgr5*^{EGFPhi} cells demonstrating functional IESC characteristics could be the result of isolation and culture conditions, it is also possible that not all of these cells are IESCs. We propose that a combinatorial approach using multiple genetic markers will provide a more robust genetic signature for the IESC and allow for a more pure population of IESCs to be isolated by FACS. As this study demonstrates, the *Sox9*^{EGFP} mouse model combined with color-gating flow cytometry provides a powerful tool to screen both known and uncharacterized cell surface markers for their capacity to uniquely mark the IESC population. The ability to use *CD24* as an IESC enrichment factor represents a new and transformative approach to IESC isolation, as well as to the study of IESC maintenance, expansion, and differentiation.

CHAPTER 4

SIGNIFICANCE AND IMPLICATIONS FOR FUTURE WORK

Defining a genetic signature for IESCs

Recent efforts in IESC biology have been intensely focused on the discovery of genetic markers that can be used to identify IESCs. *In vivo* lineage tracing was employed to validate *Lgr5* as the first IESC marker, and these results were subsequently reinforced by data demonstrating that single $Lgr5^{HIGH}$ cells fulfill requirements for 'stemness' *in vitro* (8, 112). Our research identified *Sox9* as a marker of IESCs, which was later independently confirmed by *in vivo* lineage tracing (41, 45). In addition to *Lgr5* and *Sox9*, current genes confirmed as IESC markers by *in vivo* genetic lineage tracing are: *CD133, Bmi1,* and *mTert* (87, 109, 121). *Olfm4*, *Ascl2*, and *Msi1* are also associated with IESCs based on expression patterns, but have not been confirmed with *in vivo* or *in vitro* assays (98, 139, 140).

Despite the growing number of putative and confirmed IESC markers, it remains an important caveat that no gene identified to date can be classified as an exclusive marker. All exhibit expression patterns that encompass a number of cell types within the intestinal crypts. For example, evidence suggests that *Lgr5* and *Sox9* are not exclusive markers of IESCs. $Sox9^{HIGH}$ cells have been shown to co-localize with markers of enteroendocrine cells (39). Furthermore, only $Sox9^{LOW}$- and $Lgr5^{HIGH}$-expressing cells efficiently form cryptoids *in vitro* (45, 112).

Evidence from other tissue types suggests that a combination of multiple markers may be necessary in order to accurately identify IESCs (3, 101, 104). For experimental purposes, the identification of IESCs using transgenic reporter animals and distinct gene expression levels, as demonstrated for *Sox9* and *Lgr5*, is sufficient for downstream applications (45, 112). However, transgenic expression of reporter genes cannot be exploited toward the clinical application of stem cells, an important translational goal of stem cell biology. No studies to date have isolated human intestinal epithelial cells that have demonstrated characteristics of multipotency and self-renewal. Efforts to expand human intestinal epithelium are largely hampered by a lack of a genetic signature for human IESCs that allows for FACS isolation. The identification of cell-surface markers for IESCs in animal models could be applied toward overcoming the hurdle of IESC isolation from human tissue. Our work identifies *CD24* as an enrichment factor for IESCs and suggests that this cell-surface marker might also be useful for isolation of human IESCs. As an alternative approach to identifying IESC-specific markers, genes associated with differentiated cell types could also facilitate IESC isolation through labeling and subsequent exclusion of these cells during FACS. A critical next step to enable FACS enrichment of IESCs is the identification of additional cell-surface markers that could be used in conjunction with *CD24* to establish an IESC-specific genetic signature.

***In vitro* approaches to the study of IESCs**

Until recently, the *in vivo* standard for assaying multipotency and self-renewal of putative stem cell populations in adult tissues, including the intestine, was genetic lineage tracing (125). However, this assay has obvious disadvantages in that it necessitates the

development of a transgenic mouse. Transgenic mouse development entails significant time and financial commitments, therefore an investigator must have confidence that a candidate gene will prove to be an IESC marker prior to deciding to pursue genetic lineage tracing.

The techniques facilitating culture and differentiation of single IESCs, originally published by Sato, et al. and applied in our studies of *Sox9*, are significant in that they provide an additional *in vitro* standard by which to assay multipotency and self-renewal in putative IESC populations (45, 112). Since many putative IESC markers are expressed in multiple cell types within the intestinal crypts, the ability to isolate and assay cells expressing distinct levels of a gene – such as $Sox9^{LOW}$ and $Lgr5^{HIGH}$ – allows for a level of genetic specificity for IESC biomarkers beyond the limitations of *in vivo* lineage tracing. However, several caveats regarding the currently methodology of *in vitro* IESC assays stand to be considered. One major disadvantage of the conventional Matrigel IESC culture system is that it exposes isolated stem cells to signaling from contaminating cell types as well as other IESCs. Culturing several hundred to several thousand cells in a single Matrigel patty prevents the development of a truly clonogenic cell population, and might in turn influence IESC behavior. Recent data has shown that co-culture of IESCs with Paneth cells greatly increases IESC survival, likely due to growth factors produced by Paneth cells (111). It remains to be shown whether other contaminating cells types can have positive or negative effects on normal IESC behavior *in vitro*.

Also disadvantageous to conventional Matrigel-based culturing of IESCs is that the cost of the system is prohibitive to high-throughput screening of growth factors and

pharmacologic compounds. The outcome of drug studies on IESCs *in vitro* may also be affected by contaminating cell types, as discussed above. A logical progression within the field is to establish a true clonogenic assay for the study of IESCs *in vitro*. Though no technology currently exists to facilitate this, collaborative efforts between stem cell biologists and biomedical engineers has led to the development of high-throughput assays for embryonic stem cells (6). Similar collaborations may prove advantageous to the intestinal stem cell field as well. To date, this *in vitro* system has only been used to assay isolated intestinal epithelial cells for stemness. Further technological advances might also adapt this assay to be amenable with genetic manipulation *in vitro* and provide the field with a very powerful tool for the study of molecular mechanisms in IESCs.

Mechanistic roles for *Sox* in the small intestine

Though the identification of IESC biomarkers is a vital step toward understanding intestinal stem cell biology, the field must also be concerned with identifying the mechanisms by which IESCs are capable of maintaining stemness. *Sox* proteins are particularly attractive for the study of such molecular mechanisms both as transcription factors and as mediators of stemness and differentiation in tissues throughout the developing and adult organism (76).

As discussed in previous chapters, *Sox9* has a functional role in the intestinal epithelium as demonstrated by a phenotype resulting in hyperproliferation of IESCs and progenitor cells and loss of Paneth cells in pan-epithelial knockouts (10). However, further work is required to determine the changes in normal molecular characteristics of the epithelium that result in such a phenotype. Additionally, the broad expression pattern of *Sox9* and its association with both differentiated and proliferating cell types suggests

that IESC-specific knockout is required in order to establish a definitive phenotype as relevant to stem cell biology (39, 45). The loss of Paneth cells in pan-epithelial *Sox9*-null mutants may contribute to the observed phenotype independent of *Sox9* function in IESCs, as Paneth cells have recently been shown to be critical in maintaining the IESC niche (10, 111). Existing data suggest that future studies on the role of *Sox9* in IESCs should focus on ablation of *Sox9* in specific cell types within the intestinal epithelium, including IESC- and Paneth cell-specific knockouts. Data from cell-specific knockouts could begin to outline a precise signal transduction pathway through which *Sox9* intrinsically or extrinsically confers or controls stemness in IESCs.

In addition to dissecting signaling mechanisms for *Sox* in the intestine, emerging data imply an epigenetic role for *Sox* factors in the maintanence of stemness. Recent studies have shown that stem cells contain unique chromatin modifications caused by the binding of transcription factors to *cis*-elements and have proposed that these epigenetic marks may facilitate stem cell potency (128). Specifically, *Sox2* and *Sox4* have been shown to create promoter-specific histone modifications associated with ES cells and early hematopoietic progenitors, respectively (78). This property of *Sox* factors may be of special interest in the intestinal epithelium, as *Sox4* has previously been localized to the proliferative region in the crypts (119). To date, no extensive studies regarding the role of *Sox4* in the intestinal epithelium have been carried out. However, elegant conditional and embryonic deletion of *Sox4* and the other SoxC proteins, *Sox11* and *Sox12*, has demonstrated that this subgroup of *Sox* factors is critical for stem cell survival in neural and mesenchymal tissues (11, 133). Given the central role for SoxC proteins in other stem cell populations and a known epigenetic role in pre-B cells, *Sox4* may be an

attractive target gene in the study of IESCs, especially if this mechanism of conferring potency is preserved across different tissue types.

It would also be of great interest to examine the functional relationship between *Sox4* and *Sox9*. If compensatory for one another, deletion of either *Sox4* or *Sox9* might be expected to lead to an increase in expression of the compensatory factor. However, due to the structural differences between *Sox4* and *Sox9* (leading to their classification in distinct *Sox* subgroups), they may compete for the same DNA binding site while eliciting different effects (14). This second scenario could result in appreciable phenotypic consequences. That is, an increase in *Sox4* interaction with orphaned *Sox9* binding-sites could be responsible for the phenotypes resulting from the loss of *Sox9*. Interestingly, preliminary data generated in our lab has shown that *Sox4* exhibits an expression pattern similar to *Sox9* (data not shown).

The work presented in this thesis establishes important models and reagents for the isolation of IESCs and mechanistic study of *Sox* factor contributions to stemness in the intestinal epithelium. Future studies will utilize novel conditional and temporal knockout animals to delineate phenotypes resulting from deletion of *Sox4* and *Sox9* in the intestinal epithelium as well as in specific cell types. A greater understanding of *Sox*-mediated regulation of stemness could provide the basic biological knowledge necessary to exploit tissue-specific stem cells for the treatment of injury, degenerative disease, and cancer.

Table 1. *Sox* expression in embryonic and adult gastrointestinal (epithelium) tissue.

	Expression		
Gene	**Embryonic**	**Adult**	**Reference**
Sox2	Foregut endoderm	Stomach, esophagus	(102, 131, 152)
Sox3	Posterior foregut	ND	(152)
Sox4	Pancreas	Pancreatic islets, intestine	(80, 119, 150)
Sox5	Pancreas	ND	(80, 150)
Sox6		Pancreatic islets	(57)
Sox7	Pancreas	Esophagus, stomach, jejunum, ileum, ascending colon	(80, 150)
Sox8	Pancreas	ND	(80, 150)
Sox9	Pancreas, pancreatic epithelial cords, primitive intrahepatic bile duct structures	Pancreatic islets, intestine, colon, intrahepatic bile ducts	(7, 10, 13, 39, 45, 80, 115, 116, 150)
Sox10	Pancreas	ND	(80, 150)
Sox11	Pancreas	ND	(80, 150)
Sox12	Pancreas	Pancreatic islets	(80, 150)
Sox13	Pancreas, liver	Pancreatic islets	(80, 150)
Sox15	Pancreas	ND	(80, 150)
Sox17	Pancreas, biliary primordium	Esophagus, stomach, intestine, colon	(64, 80, 119, 123, 136)
Sox18	Pancreas, liver	Stomach, intestine, colon, liver, pancreas	(80, 106)
Sox30	Pancreas	ND	(80)

Table 2. *Sox* expression in gastrointestinal pathology and disease.

Gene	Pathology	Reference
Sox2	Pancreatic intraepithelial neoplasia (up), intestinal metaplasia (down), gastric cancer (down), traceoesophageal fistula (down), ampulla of vater adenocarcinoma, Barrett's esophagus (down), esophageal squamous cell carcinoma (up)	(77, 92, 100, 102, 135)
Sox4	Colon cancer (up)	(119)
Sox6	Obesity-related insulin resistance (down)	(57)
Sox7	Gastric cancer (up, some cases), colon cancer (down, some cases)	(63)
Sox9	Colon cancer, gastric cancer (up)	(13, 158)
Sox13	Diabetes autoantigen, primary biliary cirrhosis autoantigen	(38, 62)
Sox17	Colon cancer (down)	(119)

Table 3. *Sox* expression in cell lines derived from gastrointestinal tissue.

Gene	Cell line	Reference
Sox2	MKN45$^{Hu, Stomach}$, TGBC11TKB$^{Hu, Stomach}$, KATOIII$^{Hu, Stomach}$, LoVo$^{Hu, Colon}$, CaRI$^{Hu, Colon}$, Az521$^{Hu, Stomach}$, Hep-12$^{Hu, Liver}$	(77, 92, 131, 155)
Sox7	BxPC3$^{Hu, Pancreas}$, PSN1$^{Hu, Pancreas}$, Hs766T$^{Hu, Pancreas}$, MKN45$^{Hu, Stomach}$, TE2$^{Hu, Esophagus}$, TE3$^{Hu, Esophagus}$, TE4$^{Hu, Esophagus}$, TE5$^{Hu, Esophagus}$, TE7$^{Hu, Esophagus}$, TE8$^{Hu, Esophagus}$, TE11$^{Hu, Esophagus}$, TE12$^{Hu, Esophagus}$, TE13$^{Hu, Esophagus}$	(63)
Sox9	HT29Cl.16E$^{Hu, Colon}$, LS174T$^{Hu, Colon}$, DLD1$^{Hu, Colon}$, HTC116$^{Hu, Colon}$, SW480$^{Hu, Colon}$, TC7$^{Hu, Colon}$	(13)
Sox17	SW480$^{Hu, Colon}$	(64)
Sox18	SW480$^{Hu, Colon}$, TMK1$^{Hu, Stomach}$, MKN45$^{Hu, Stomach}$, PANC1$^{Hu, Pancreas}$, BxPC3$^{Hu, Pancreas}$, AsPC1$^{Hu, Pancreas}$, PSN1$^{Hu, Pancreas}$, Hs700T$^{Hu, Pancreas}$, Hs766T$^{Hu, Pancreas}$, MIA PaCa2$^{Hu, Pancreas}$	(106)

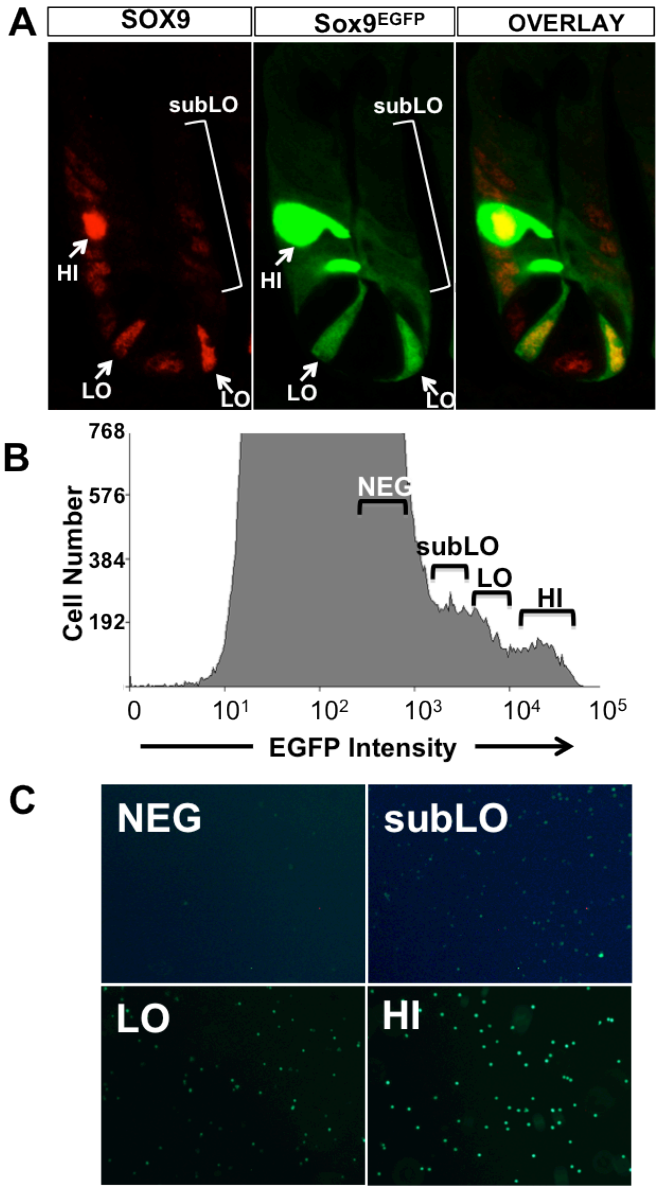
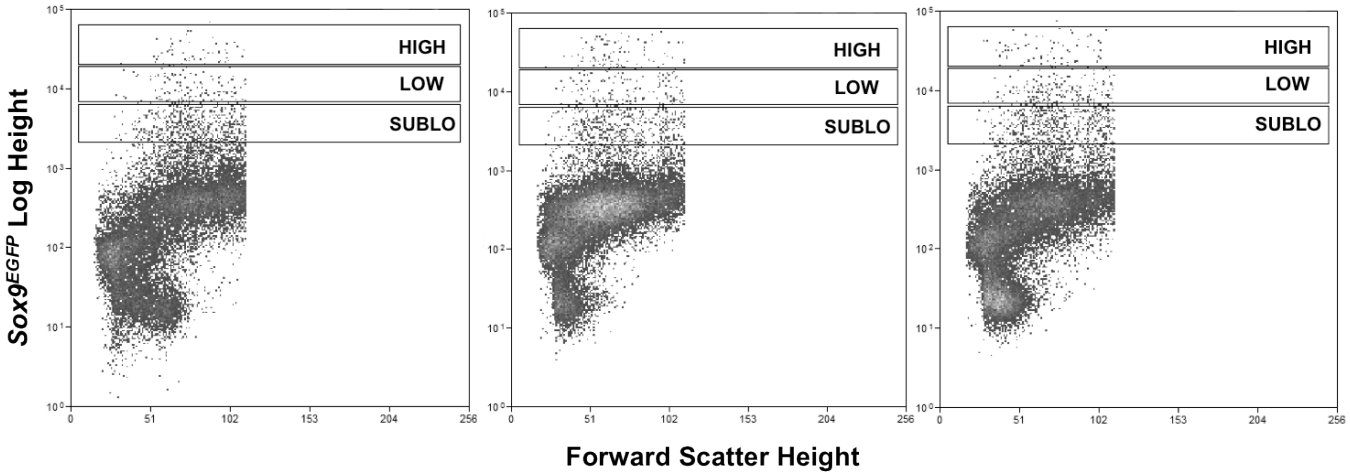


Figure 1. *Sox9*^{EGFP} is expressed at 3 different levels in the small intestinal epithelium. *Sox9EGFP* is expressed at 3 different levels in the small intestinal epithelium. **A**: *Sox9EGFP* is expressed at variable levels, "HI," "LO," and "subLO," in the crypts of the small intestine. HI levels are associated with postmitotic enteroendocrine cells, LO levels are associated with crypt-base columnar cells, "subLO" levels are associated with the transit-amplifying region of the crypt, and NEG levels are *Sox9EGFP* negative. Images represent 1,260X original magnification. **B**: flow cytometric analysis indicates distinct *Sox9EGFP* expression levels. Gate parameters used to sort each population are indicated above each region of the histogram. **C**: postsort analysis indicates that single *Sox9EGFP*-expressing cells have been isolated based on enhanced green fluorescent protein (EGFP) status. Images represent 200X original magnification. The image exposures for *Sox9EGFPneg* and *Sox9EGFsubLo* panels were doubled to produce images that would allow visualization of EGFP expression in these 2 populations.

A



B

Population	Experiment 1	Experiment 2	Experiment 3	Average
Sox9^{EGFPhi}	0.25%	0.35%	0.32%	0.31% ± 0.05%
Sox9^{EGFPlo}	0.62%	0.59%	0.85%	0.69% ± 0.14%
Sox9^{EGFPsubLo}	2.96%	1.72%	2.58%	2.42% ± 0.64%

* Percentages represent the numbers of cells within each gate

Figure 2. *Sox9^{EGFP}* populations are reproducibly separated by FACS. A: Three separate FACS experiments performed on three different animals were analyzed to demonstrate that gating protocols for isolating *Sox9EGFPsubLo*, *Sox9EGFPlo*, and *Sox9EGFPhi* are reproducible based on histograms analyzing EGFP fluorescence. **B:** Numerical analysis of these data shows that the percent of sorted cells falling in each gate is consistent between experiments, demonstrating reproducibility in the separation of *Sox9EGFP* populations. Cut-off of the events on the right side of the histogram is due to gating parameters to exclude potential doublets.

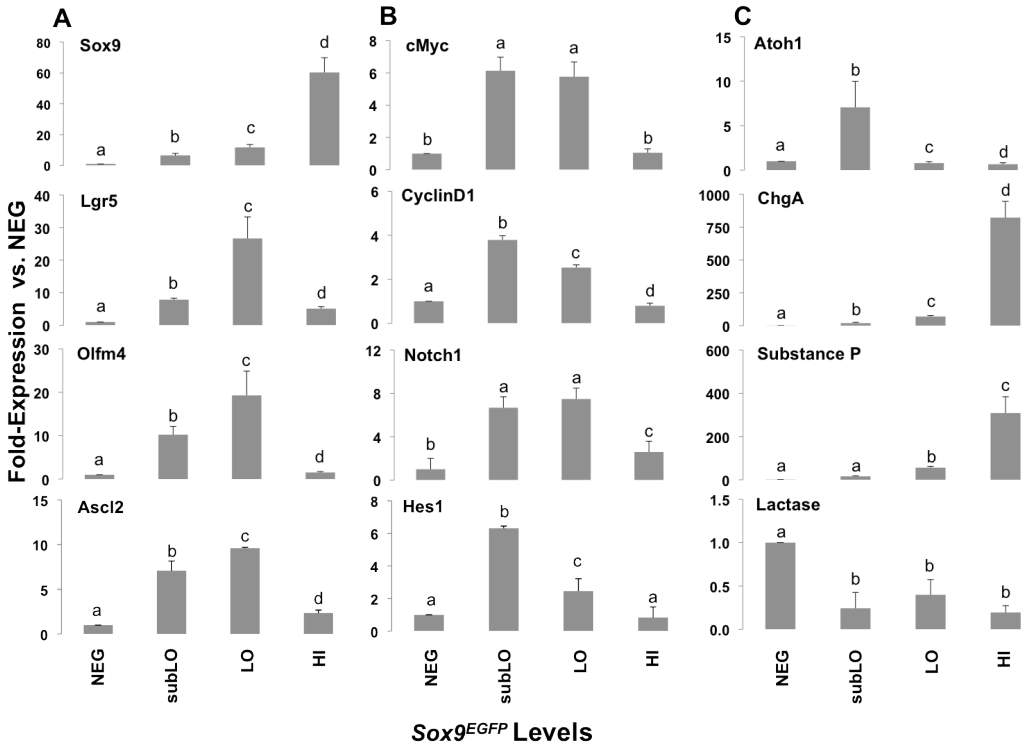


Figure 3. Gene expression analysis demonstrates isolation and enrichment of intestinal epithelial stem cells (IESCs), progenitors, and enteroendocrine cells by use of fluorescence-activated cell sorting (FACS). Semiquantitative RT-PCR conducted on FACSed NEG, subLO, LO, and HI cells demonstrates enrichment of IESC stem cell biomarkers in the LO populations **A:** active *Wnt/β-catenin* and *Notch1* signaling in subLO and LO populations **B:** and enrichment of committed secretory progenitors in the subLO population and enteroendocrine cells in the HI population. **C:** Elevated lactase expression in the NEG population indicates this population is enriched for enterocytes. All data points represent means ± SE from 3 independent experiments; statistical analysis was by ANOVA and post hoc 2 sample *t*-tests were then performed. A P value < 0.05 is considered statistically significant. Different letters above each bar represent data points that are statistically different from each other.

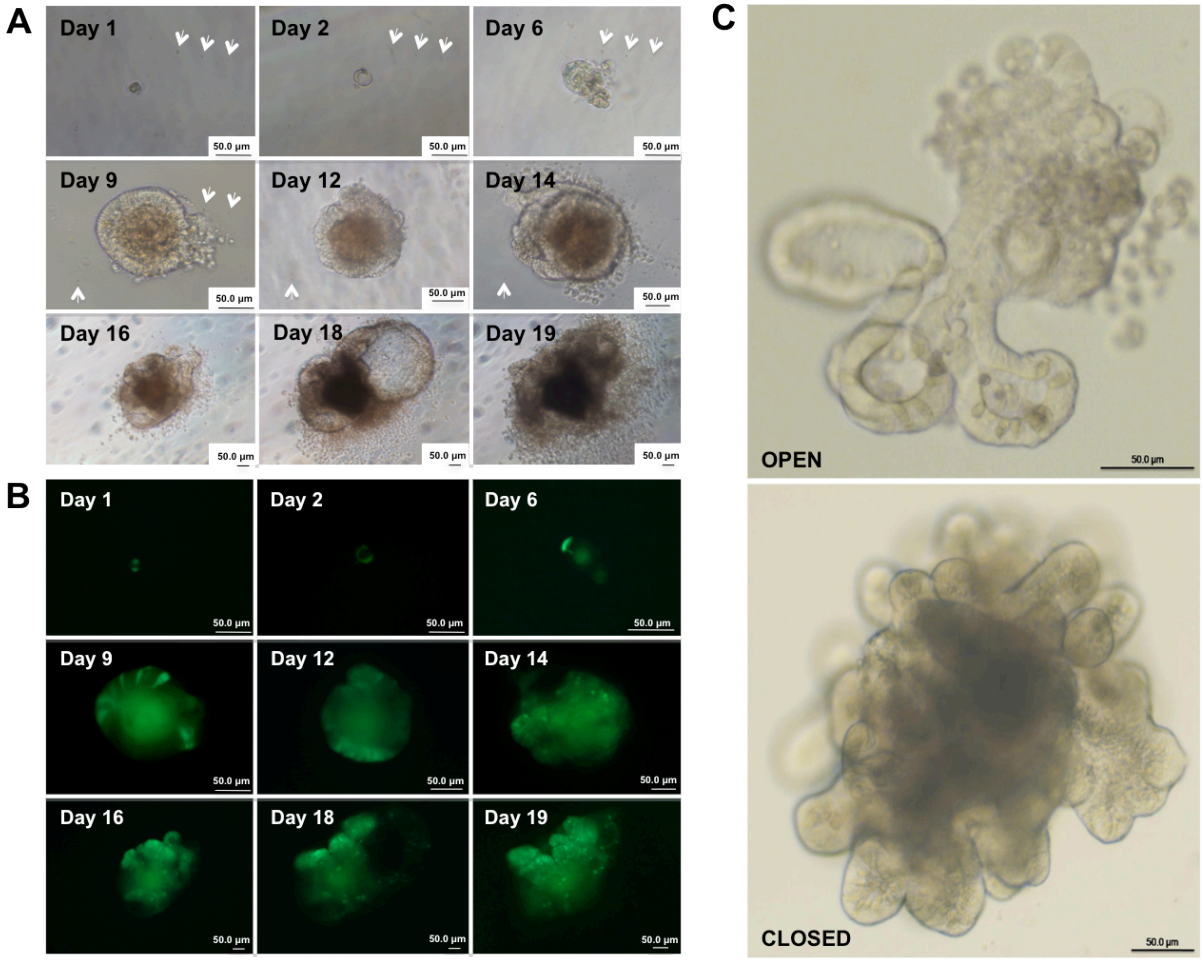


Figure 4. Single *Sox9EGFPlo* cells form complex organoids in vitro. A: single cells develop into large aggregates with crypt and villus-like structures over 7–10 days. The central pseudolumens of the single-cell-derived organoids expel apoptotic cells resulting from the sloughing of terminally differentiated epithelial cells. Arrows mark defects in culture plastic used as "landmarks" to track individual cells through early developmental phases into formed organoids. At day 9 the organoid grew large enough to obscure the tracking defects; therefore, additional tracking marks were identified (white arrows). **B**: epifluorescent images of organoids depicted in A. Note that the green pseudolumen is nonspecific autofluorescence and not EGFP fluorescence. Differential *Sox9EGFP* expression patterning remains evident throughout expansion in vitro. **C**: organoids form both open and closed morphologies, based on developmental positioning of the apoptotic pseudolumen. Apoptotic cells slough openly into Matrigel in the "open" morphology and are seen as a dense, dark region in the "closed" morphology. Ages of organoids are 12 days postsort (top) and 15 days postsort (bottom). Organoids are magnified 300X (top) and 200X (bottom).

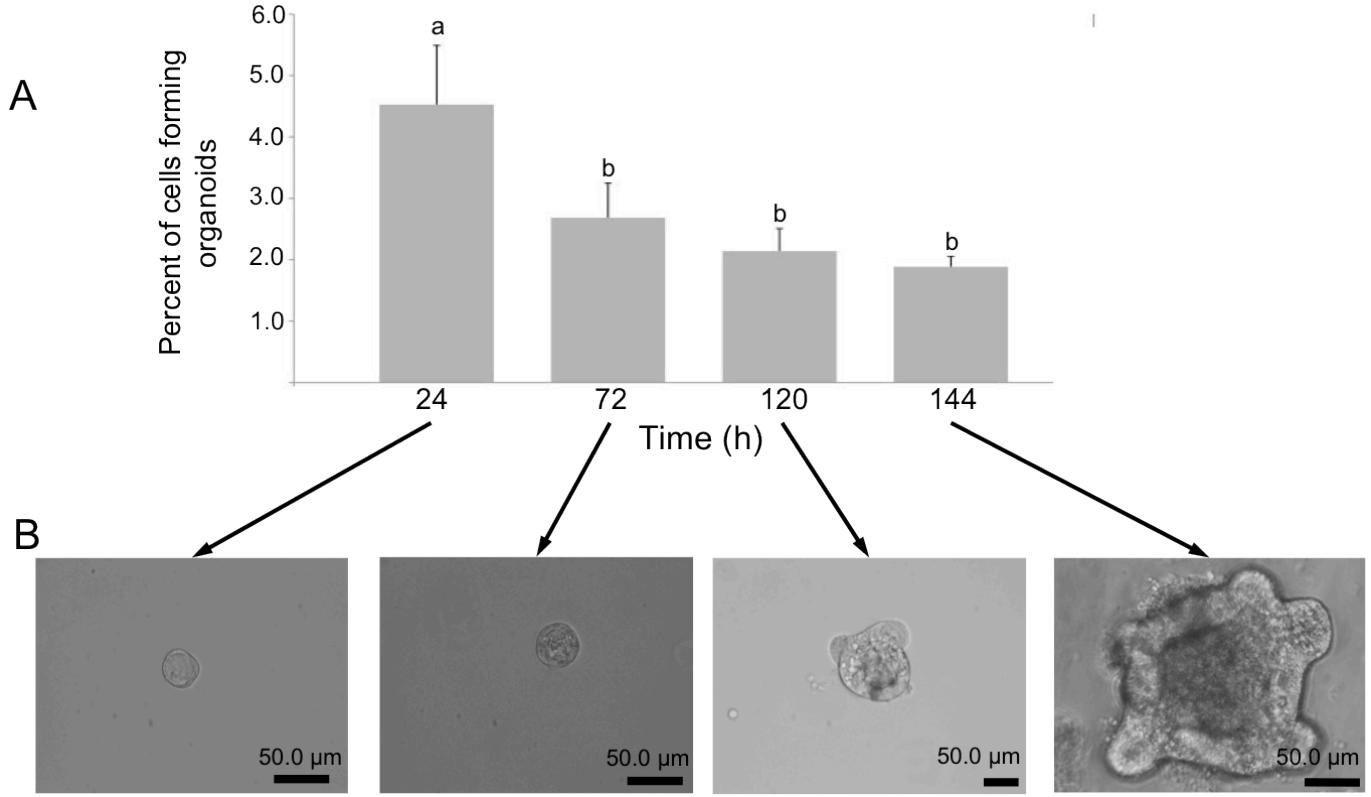


Figure 5. Organoid survival is quantified by identification of unique morphology. A: Organoids were quantified in culture at 24-hour intervals to determine survival rates. **A & B:** Following an initial drop, organoid survival at 120-144 hours leveled off at the same time crypt- like budding was observed. **B:** Organoids were scored as surviving if observed to have a characteristic morphology, marked by cellular organization in a round structure surrounding an apparently empty (24 hours) or apoptotic (120 hours) pseudo-lumen. All data points represent means ± SE from three independent experiments; statistical analysis was by ANOVA and post-hoc two sample *t*-tests were then performed. A p-value < 0.05 is considered statistically significant. The different letters above each bar represent data points that are statistically different from each other.

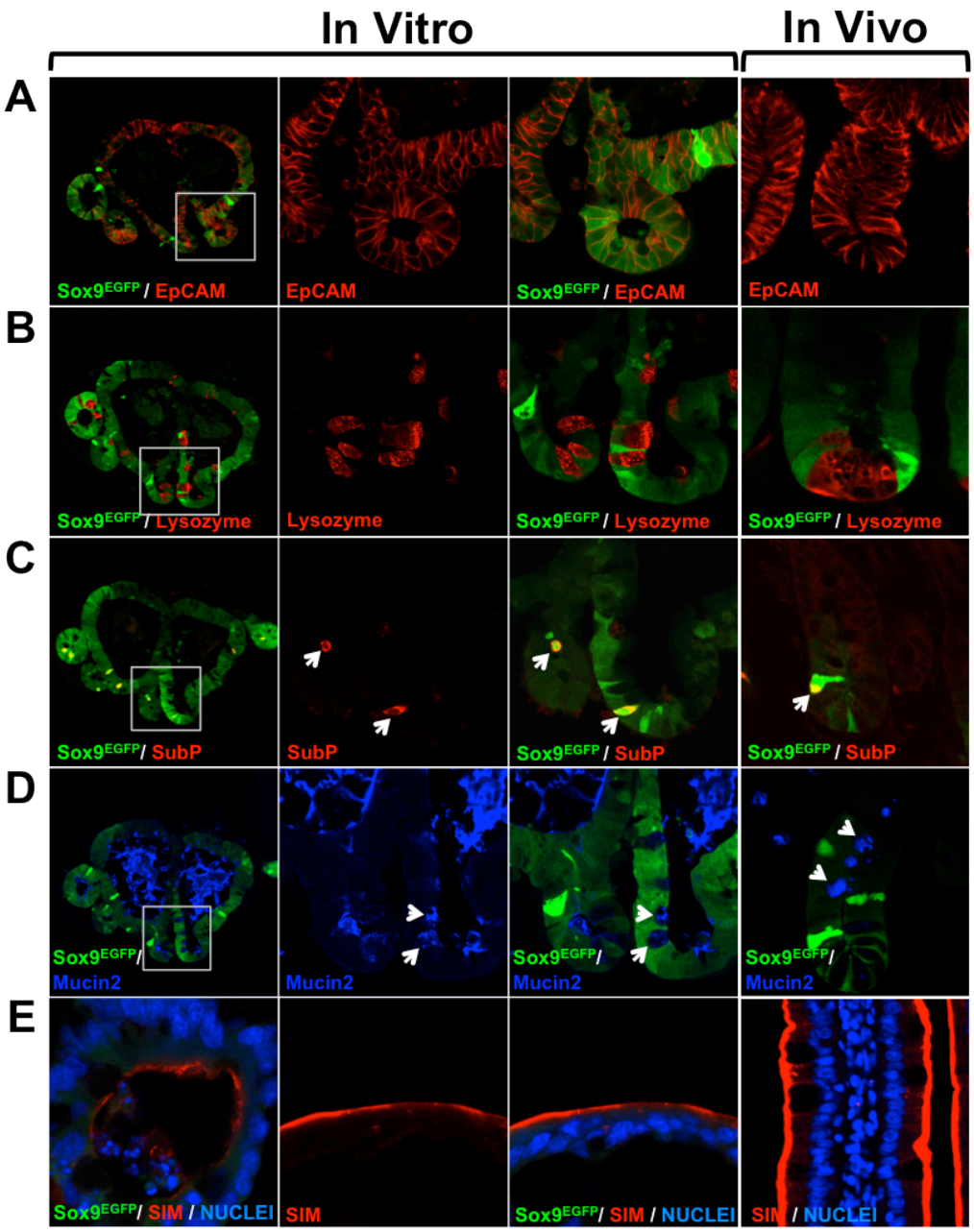


Figure 6. *Sox9EGFPlo* cells are capable of differentiation into all postmitotic cell types associated with the intestinal epithelium. A: epithelial cell adhesion molecule (*EpCAM*) staining indicates organoids cells are epithelial. **B:** lysozyme marks Paneth cells near the base of crypt-like structures. **C:** enteroendocrine cells are labeled with *Substance P* and also restricted to the base of crypt-like structures in the organoids. **D:** Mucin2 demonstrates the goblet cell phenotype in crypt-like structures. Secreted Mucin2 in the pseudolumen suggests that goblet cells are correctly polarized. **E:** sucrase isolmaltase (SIM) labels the apical ends of enterocytes lining the pseudolumen in the organoids. White boxes in the far left images (100X original magnification) represent areas magnified in the center 2 columns (800X original magnification). White arrows point to representative cells. Associated expression patterns *in vitro* correlate with those observed *in vivo* (far right). The organoid represented was cultured for 12 days.

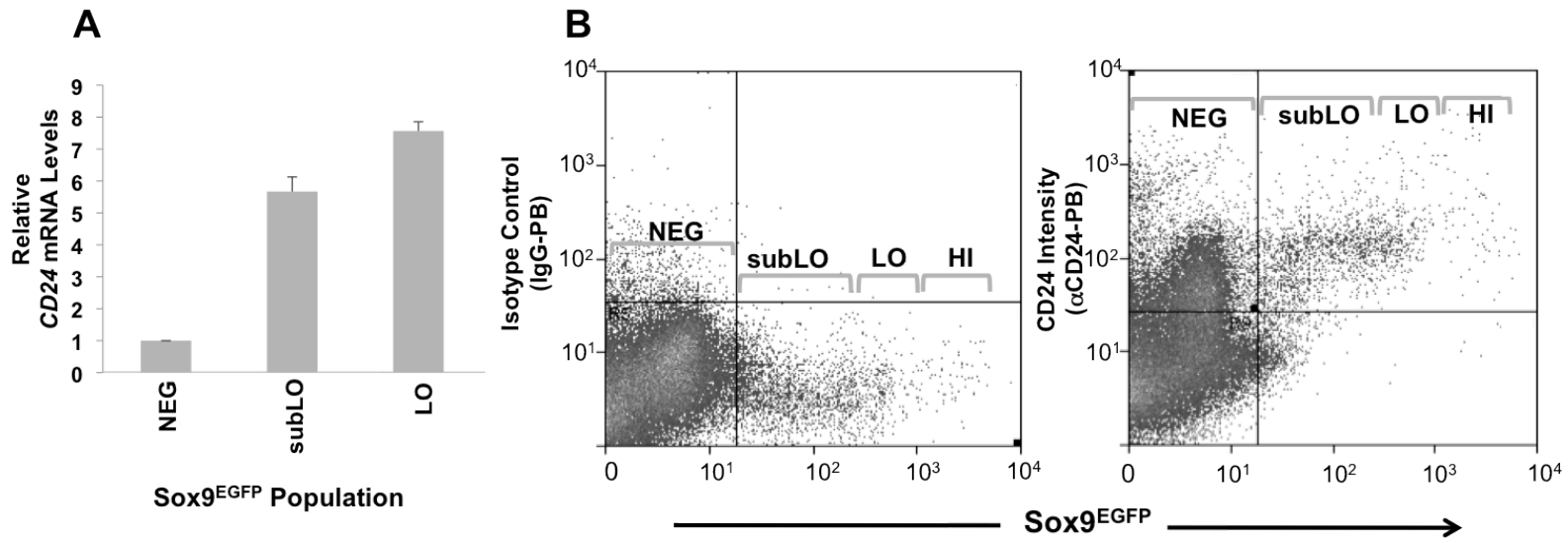


Figure 7. Small intestine epithelial stem/progenitor cells express CD24 mRNA and protein. A: Semiquantitative RT-PCR (sqRT-PCR) demonstrates that *CD24* is expressed 5.5- to 7.2-fold higher in the subLO and LO cells compared with *Sox9EGFP*-negative cells. All data points represent means ± SE from 3 independent experiments; statistical analysis was by ANOVA, and post hoc 2-sample *t*-tests were then performed. A P value < 0.05 is considered statistically significant. Different letters above each bar represent data points that are statistically different from each other. **B:** flow cytometric analysis indicates that nearly all *Sox9EGFP*-positive cells express *CD24* protein. Left: Pacific blue-conjugated isotype antibody control. Right: CD24-Pacific blue conjugated antibody. Gray brackets in the histograms represent the FACS gating parameters.

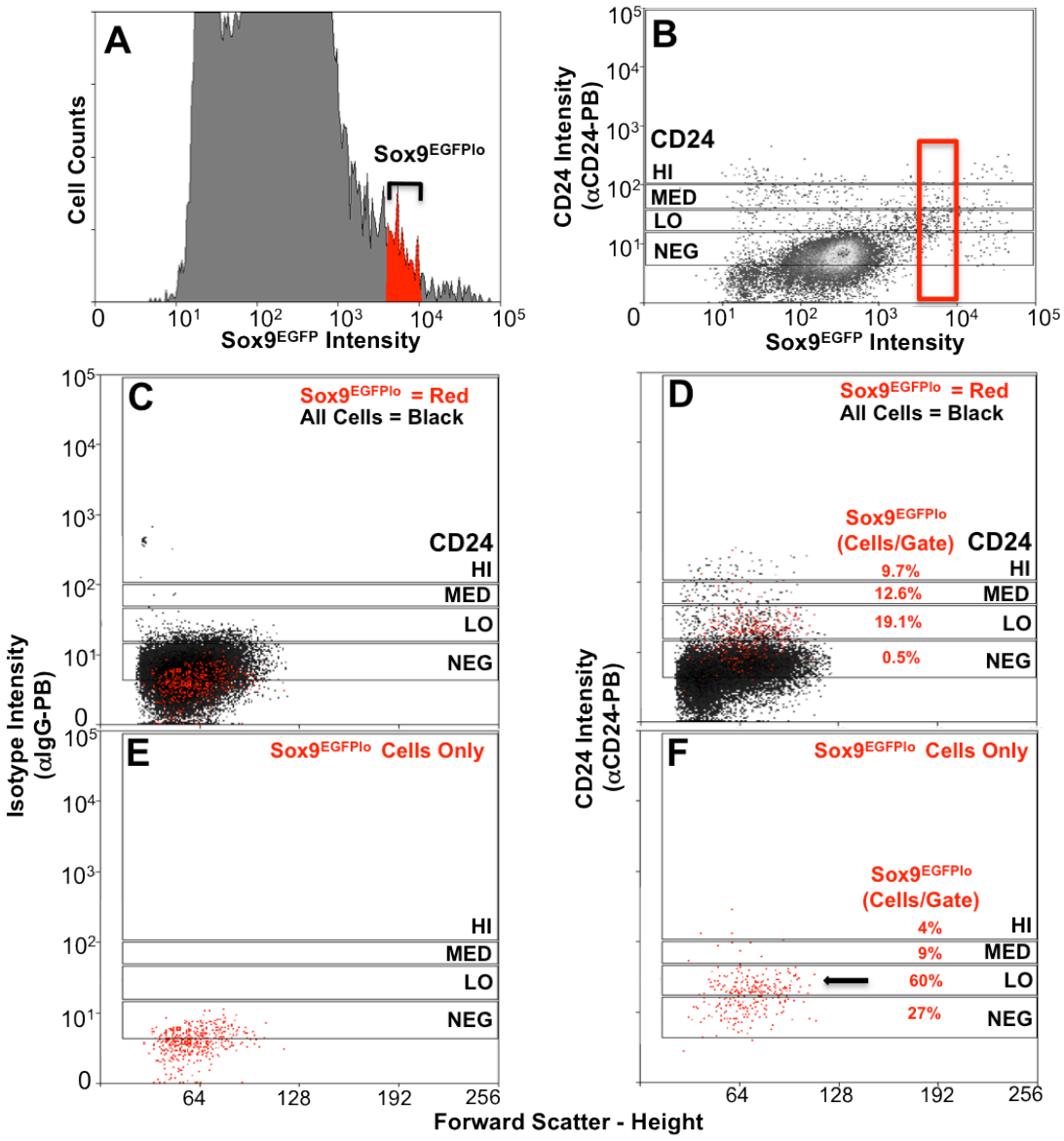


Figure 8. Flow cytometric analysis shows that a majority *Sox9EGFPlo* cells fall within CD24lo expression pattern. A: *Sox9EGFP* expression of dissociated small intestine epithelial cells on a univariate EGFP histogram. Black brackets indicate gate parameters. Color gating (**red**) allows visualization of *Sox9EGFPlo* population in histograms A-F. **B:** *Sox9EGFP*/CD24 bivariate histogram used to define the 4 CD24 gate parameters [negative (NEG), low (LO), medium (MED), and high (HI)]. **C** and **E**: IgG control antibody indicates there is no significant nonspecific staining. **D** and **F:** *CD24-PB* (Pacific blue) antibody labels 73% of *Sox9EGFPlo* cells. **D:** percentages represent the number of *Sox9EGFPlo*-expressing cells in each gate. Mean percentage ± SE for each population are as follows: NEG, 1.67 ± 0.25%; LO, 25.63 ± 0.95%; MED, 12.23 ± 1.48%; HI, 9.33 ± 0.91%. **E** and **F:** just *Sox9EGFPlo* cells color backgated onto the *CD24*/FSC histogram to highlight their distribution on the histogram. **F:** percentages represent the relative number of all *Sox9EGFPlo* cells in each gate. Mean percentage ± SE for each population are as follows: NEG, 50.7 ± 5.37%; LO, 42.47 ± 4.34%; MED, 10.23 ± 1.16%; HI, 0.50 ± 0.00%.

A

CD24 GATE	% SOX9^{EGFPlo} (per gate)	Fold Enrichment of Sox9^{EGFPlo} (vs. un-sorted)	% Organoids
HI	9.7 %	17.6X	0
MED	12.6%	24.0X	0.55 ± 0.08%
LO	19.1%	25.1X	0.19 ± 0.07%
NEG	0.5%	0.4X	0
UN-SORTED	1.1%	1X	0

B

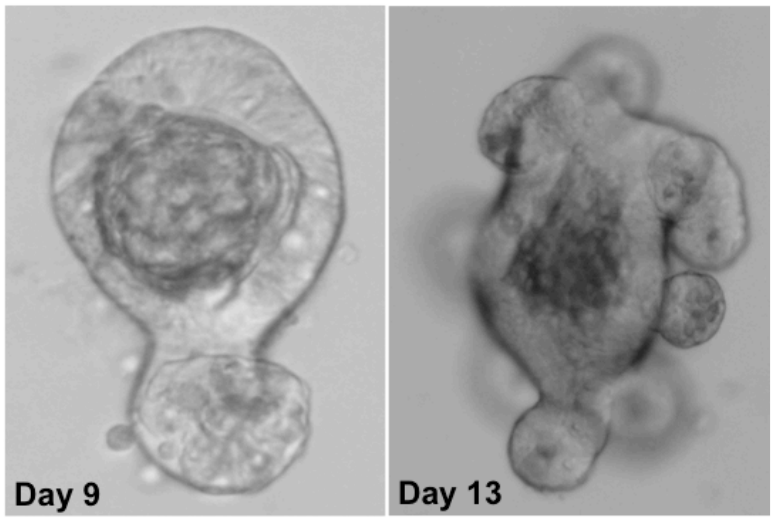


Figure 9. CD24 can be used as a biomarker to isolate and enrich for single cells capable of producing sustainable, differentiated organoids. A: table representing the enrichment of putative IESCs by using *CD24* expression levels as sort criterion. Percent *Sox9EGFPlo* cells/gate ± number of *Sox9EGFPlo* cells/total number of cells per gate. Fold enrichment, percent *Sox9EGFPlo* cells per gate/percent *Sox9EGFPlo* cells in all sorted cells. *CD24lo* sort parameters demonstrate the highest potential for enrichment of IESCs. There is a 25.1X enrichment for *Sox9EGFPlo* cells in the *CD24lo* sort parameters compared with ungated cells. Organoid generating IESCs can be isolated only with the *CD24lo* (5 organoids in every 1,000 cells) and *CD24med* (1 organoid in every 1,000 cells) sort parameters. **B:** representative organoids derived from single *CD24lo*-expressing cells (*left*, *day 9*; *right*, *day 13*).

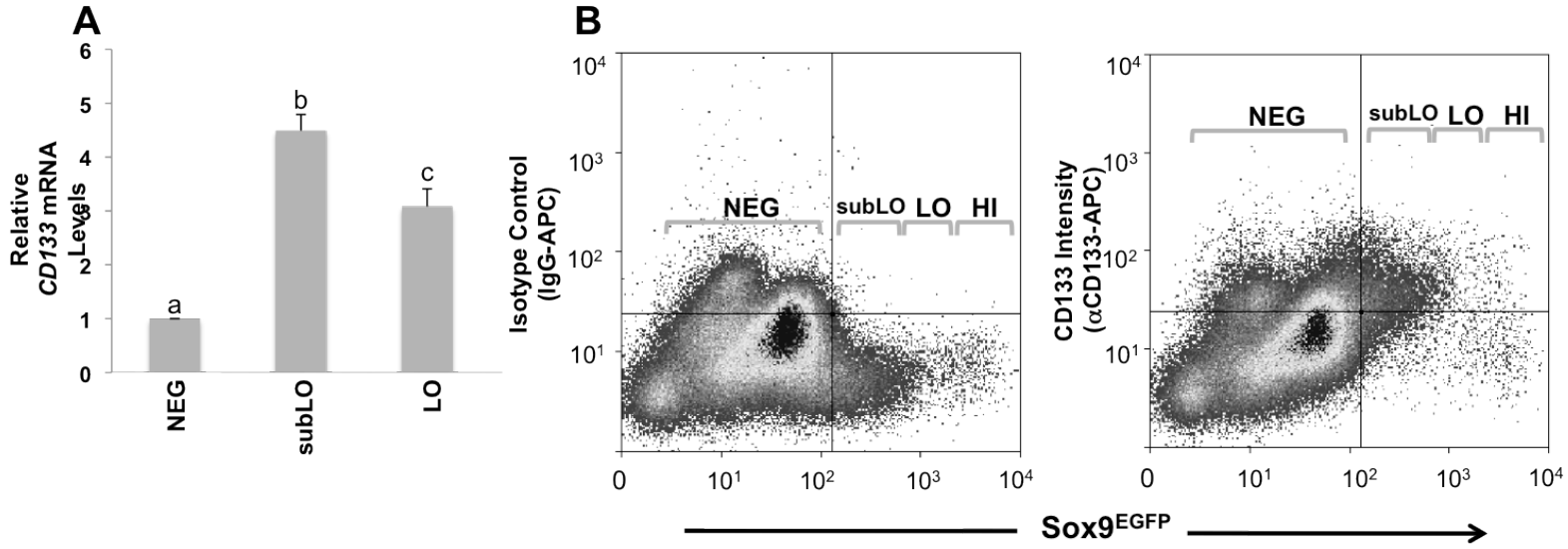


Figure 10. Small intestine epithelial stem/progenitor cells express CD133 mRNA and protein – A: sqRT-PCR demonstrates *CD133* is expressed 4.4-fold and 3.1-fold higher in the 'subLO' and 'LO' cells respectively when compared to *Sox9EGFP* negative cells. All data points represent means ± S.E.M. from three independent experiments; statistical analysis was by ANOVA and post-hoc two sample *t*-tests were then performed. A p-value < 0.05 is considered statistically significant. The different letters above each bar represent data points that are statistically different from each other. **B:** Flowcytometric analysis indicates that most *Sox9EGFP*-positive cells (94.8%) express *CD133* protein. (*left panel)* APC-conjugated isotype antibody control. αCD133-APC conjugated antibody. (*right panel)* Grey brackets in the histograms represent the FACS gating parameters.

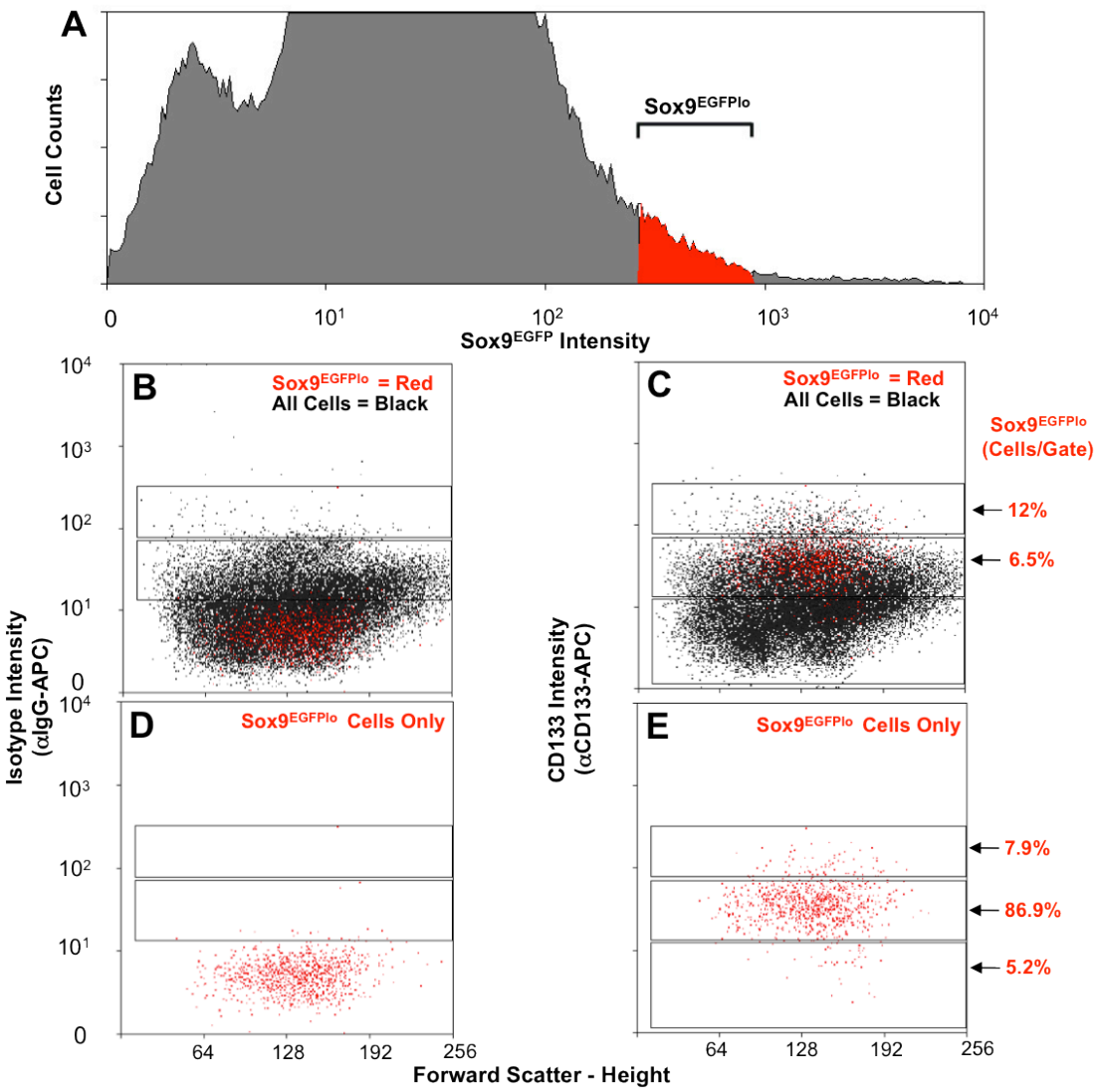


Figure 11. Using *CD133* as an IESC FACS-enrichment marker is less efficient than *CD24* – A: *Sox9EGFP* expression of dissociated small intestine epithelial cells on a univariate EGFP histogram. Black brackets indicate gate parameters. Color gating *(red)* allows visualization of *Sox9EGFPlo* population in histograms A-E. **B** and **D:** IgG control antibody indicates there is no significant non- specific staining. **C** and **E:** α*CD133-APC* antibody labels 94.8% of *Sox9EGFPlo* cells. **D & E:** Represent just *Sox9EGFPlo* cells color back-gated onto the *CD133*/FSC histogram to highlight their distribution on the histogram. **E:** Percentages represent the relative number of all *Sox9EGFPlo* cells in each gate. **C & E:** Note that although 94.8% of *Sox9EGFPlo* cells express CD133, the shift of the *Sox9EGFPlo* population on the CD133 axis is too small to sufficiently separate the *Sox9EGFPlo* cells away from non-*Sox9EGFPlo* cells. The upper-gates in each histogram represent the region of CD133-expressing cells that fall above the isotype control negative population. The lower-gates in each histogram represent the region of CD133-expressing *Sox9EGFPlo* cells that fall above the isotype control negative population.

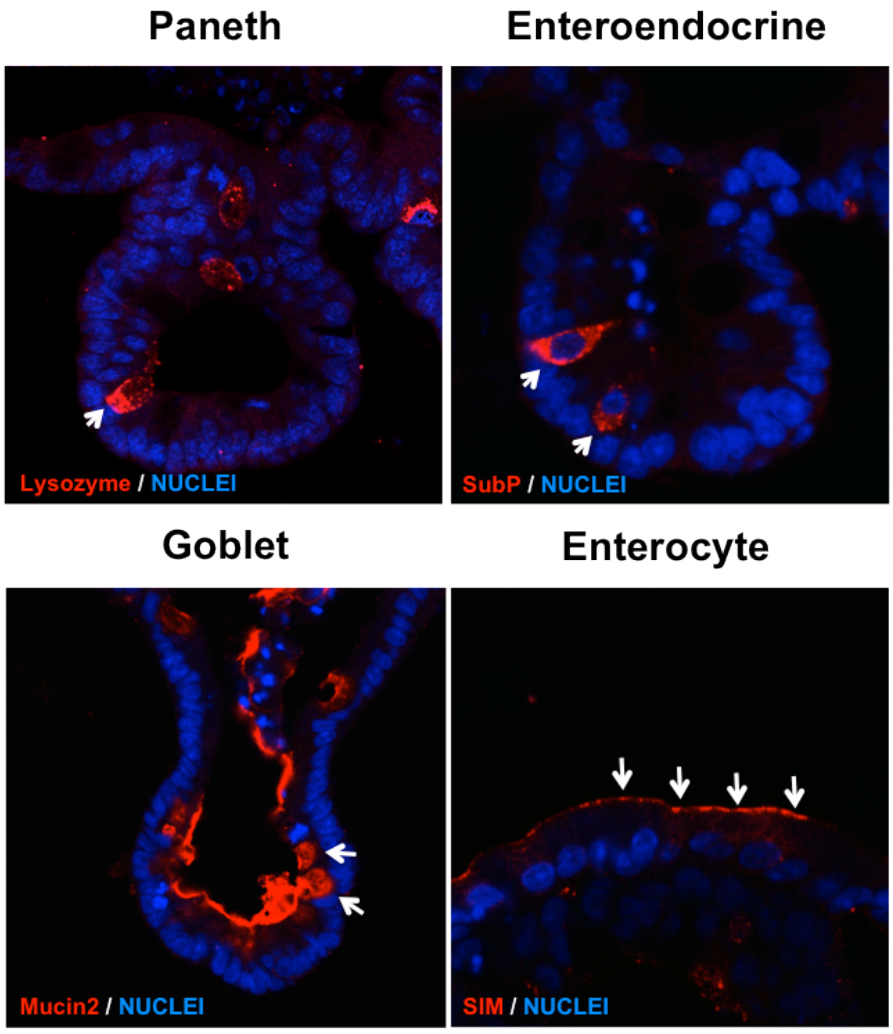


Figure 12. *CD24*-expressing IESCs are multipotent – 14 day organoids derived from CD24-FACSed IESCs were assessed for the presence of the four differentiated cells lineages of the small intestine epithelium by immunostaining. **A:** Paneth cells are marked by *Lysozyme* expression (*red*). **B:** Enteroendocrine cells are marked by *Substance P* (*red*). **C:** Goblet cells are marked by *Mucin2* staining (*red*). **D:** Absorptive enterocytes are marked by the brush border enzyme *Sucrase Isomaltase* (SIM – *red*). All images are 1260X original magnification.

REFERENCES

1. Akiyama H, Kim JE, Nakashima K, Balmes G, Iwai N, Deng JM, Zhang Z, Martin JF, Behringer RR, Nakamura T, and de Crombrugghe B. Osteo-chondroprogenitor cells are derived from Sox9 expressing precursors. *Proc Natl Acad Sci U S A* 102: 14665-14670, 2005.

2. Akiyama H, Lyons JP, Mori-Akiyama Y, Yang X, Zhang R, Zhang Z, Deng JM, Taketo MM, Nakamura T, Behringer RR, McCrea PD, and de Crombrugghe B. Interactions between Sox9 and beta-catenin control chondrocyte differentiation. *Genes Dev* 18: 1072-1087, 2004.

3. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, and Clarke MF. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* 100: 3983-3988, 2003.

4. Alison MR, Golding M, Sarraf CE, Edwards RJ, and Lalani EN. Liver damage in the rat induces hepatocyte stem cells from biliary epithelial cells. *Gastroenterology* 110: 1182-1190, 1996.

5. Ambrosetti DC, Basilico C, and 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.

6. Anderson DG, Levenberg S, and Langer R. Nanoliter-scale synthesis of arrayed biomaterials and application to human embryonic stem cells. *Nat Biotechnol* 22: 863-866, 2004.

7. Antoniou A, Raynaud P, Cordi S, Zong Y, Tronche F, Stanger BZ, Jacquemin P, Pierreux CE, Clotman F, and Lemaigre FP. Intrahepatic bile ducts develop according to a new mode of tubulogenesis regulated by the transcription factor SOX9. *Gastroenterology* 136: 2325-2333, 2009.

8. Barker N, van Es JH, Kuipers J, Kujala P, van den Born M, Cozijnsen M, Haegebarth A, Korving J, Begthel H, Peters PJ, and Clevers H. Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* 449: 1003-1007, 2007.

9. Bass AJ, Watanabe H, Mermel CH, Yu S, Perner S, Verhaak RG, Kim SY, Wardwell L, Tamayo P, Gat-Viks I, Ramos AH, Woo MS, Weir BA, Getz G, Beroukhim R, O'Kelly M, Dutt A, Rozenblatt-Rosen O, Dziunycz P, Komisarof J, Chirieac LR, Lafargue CJ, Scheble V, Wilbertz T, Ma C, Rao S, Nakagawa H, Stairs DB, Lin L, Giordano TJ, Wagner P, Minna JD, Gazdar AF, Zhu CQ, Brose MS, Cecconello I, Jr UR, Marie SK, Dahl O, Shivdasani RA, Tsao MS, Rubin MA, Wong KK, Regev A, Hahn WC, Beer DG, Rustgi AK, and Meyerson M. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Nat Genet* 41: 1238-1242, 2009.

10. Bastide P, Darido C, Pannequin J, Kist R, Robine S, Marty-Double C, Bibeau F, Scherer G, Joubert D, Hollande F, Blache P, and Jay P. Sox9 regulates cell proliferation and is required for Paneth cell differentiation in the intestinal epithelium. *J Cell Biol* 178: 635-648, 2007.

11. Bhattaram P, Penzo-Mendez A, Sock E, Colmenares C, Kaneko KJ, Vassilev A, Depamphilis ML, Wegner M, and Lefebvre V. Organogenesis relies on SoxC transcription factors for the survival of neural and mesenchymal progenitors. *Nat Commun* 1: 9, 2010.

12. Bjerknes M, and Cheng H. The stem-cell zone of the small intestinal epithelium. III. Evidence from columnar, enteroendocrine, and mucous cells in the adult mouse. *Am J Anat* 160: 77-91, 1981.

13. Blache P, van de Wetering M, Duluc I, Domon C, Berta P, Freund JN, Clevers H, and Jay P. 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.

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

15. Brennand K, and Melton D. Slow and steady is the key to beta-cell replication. *J Cell Mol Med* 13: 472-487, 2009.

16. Brockenbrough JS, Weir GC, and Bonner-Weir S. Discordance of exocrine and endocrine growth after 90% pancreatectomy in rats. *Diabetes* 37: 232-236, 1988.
17. Bucher NL, and Swaffield MN. The Rate of Incorporation of Labeled Thymidine into the Deoxyribonucleic Acid of Regenerating Rat Liver in Relation to the Amount of Liver Excised. *Cancer Res* 24: 1611-1625, 1964.

18. Bucher NM, RA. *Regeneration of liver and kidney*. Boston, MA: Little, Brown, 1971.

19. Chen X, Qin R, Liu B, Ma Y, Su Y, Yang CS, Glickman JN, Odze RD, and Shaheen NJ. Multilayered epithelium in a rat model and human Barrett's esophagus: similar expression patterns of transcription factors and differentiation markers. *BMC Gastroenterol* 8: 1, 2008.

20. 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.

21. 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.

22. Cheng H, and Leblond CP. 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.

23. Cheng H, and Leblond CP. 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.

24. Cheng H, and Leblond CP. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. V. Unitarian Theory of the origin of the four epithelial cell types. *Am J Anat* 141: 537-561, 1974.

25. Cheng H, Merzel J, and Leblond CP. Renewal of Paneth cells in the small intestine of the mouse. *Am J Anat* 126: 507-525, 1969.

26. Clarke AR. Wnt signalling in the mouse intestine. *Oncogene* 25: 7512-7521, 2006.

27. Collignon J, Sockanathan S, Hacker A, Cohen-Tannoudji M, Norris D, Rastan S, Stevanovic M, Goodfellow PN, and Lovell-Badge R. A comparison of the properties of Sox-3 with Sry and two related genes, Sox-1 and Sox-2. *Development* 122: 509-520, 1996.

28. Connor F, Cary PD, Read CM, Preston NS, Driscoll PC, Denny P, Crane-Robinson C, and Ashworth A. DNA binding and bending properties of the post-meiotically expressed Sry-related protein Sox-5. *Nucleic Acids Res* 22: 3339-3346, 1994.

29. Crane-Robinson C, Read CM, Cary PD, Driscoll PC, Dragan AI, and Privalov PL. The energetics of HMG box interactions with DNA. Thermodynamic description of the box from mouse Sox-5. *J Mol Biol* 281: 705-717, 1998.

30. Croagh D, Phillips WA, Redvers R, Thomas RJ, and Kaur P. Identification of candidate murine esophageal stem cells using a combination of cell kinetic studies and cell surface markers. *Stem Cells* 25: 313-318, 2007.

31. Croagh D, Thomas RJ, Phillips WA, and Kaur P. Esophageal stem cells--a review of their identification and characterization. *Stem Cell Rev* 4: 261-268, 2008.

32. Dekaney CM, Fong JJ, Rigby RJ, Lund PK, Henning SJ, and Helmrath MA. Expansion of intestinal stem cells associated with long-term adaptation following ileocecal resection in mice. *Am J Physiol Gastrointest Liver Physiol* 293: G1013-1022, 2007.

33. Dekaney CM, Rodriguez JM, Graul MC, and Henning SJ. Isolation and characterization of a putative intestinal stem cell fraction from mouse jejunum. *Gastroenterology* 129: 1567-1580, 2005.

34. Denny P, Swift S, Connor F, and Ashworth A. An SRY-related gene expressed during spermatogenesis in the mouse encodes a sequence-specific DNA-binding protein. *EMBO J* 11: 3705-3712, 1992.

35. Dor Y, Brown J, Martinez OI, and Melton DA. Adult pancreatic beta-cells are formed by self-duplication rather than stem-cell differentiation. *Nature* 429: 41-46, 2004.

36. Duncan AW, Dorrell C, and Grompe M. Stem cells and liver regeneration. *Gastroenterology* 137: 466-481, 2009.

37. Ferrari S, Harley VR, Pontiggia A, Goodfellow PN, Lovell-Badge R, and Bianchi ME. SRY, like HMG1, recognizes sharp angles in DNA. *EMBO J* 11: 4497-4506, 1992.

38. Fida S, Myers MA, Whittingham S, Rowley MJ, Ozaki S, and Mackay IR. Autoantibodies to the transcriptional factor SOX13 in primary biliary cirrhosis compared with other diseases. *J Autoimmun* 19: 251-257, 2002.

39. Formeister EJ, Sionas AL, Lorance DK, Barkley CL, Lee GH, and Magness ST. 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.

40. Foster JW, Dominguez-Steglich MA, Guioli S, Kwok C, Weller PA, Stevanovic M, Weissenbach J, Mansour S, Young ID, Goodfellow PN, and et al. Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. *Nature* 372: 525-530, 1994.

41. Furuyama K, Kawaguchi Y, Akiyama H, Horiguchi M, Kodama S, Kuhara T, Hosokawa S, Elbahrawy A, Soeda T, Koizumi M, Masui T, Kawaguchi M, Takaori K, Doi R, Nishi E, Kakinoki R, Deng JM, Behringer RR, Nakamura T, and Uemoto S. Continuous cell supply from a Sox9-expressing progenitor zone in adult liver, exocrine pancreas and intestine. *Nat Genet* 43: 34-41, 2011.

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

43. Gong S, Zheng C, Doughty ML, Losos K, Didkovsky N, Schambra UB, Nowak NJ, Joyner A, Leblanc G, Hatten ME, and Heintz N. A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* 425: 917-925, 2003.

44. Goodfellow PN, and Lovell-Badge R. SRY and sex determination in mammals. *Annu Rev Genet* 27: 71-92, 1993.

45. Gracz AD, Ramalingam S, and Magness ST. 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.

46. Gubbay J, Collignon J, Koopman P, Capel B, Economou A, Munsterberg A, Vivian N, Goodfellow P, and Lovell-Badge R. A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature* 346: 245-250, 1990.

47. Gulati AS, Ochsner SA, and Henning SJ. Molecular properties of side population-sorted cells from mouse small intestine. *Am J Physiol Gastrointest Liver Physiol* 294: G286-294, 2008.

48. Hacker A, Capel B, Goodfellow P, and Lovell-Badge R. Expression of Sry, the mouse sex determining gene. *Development* 121: 1603-1614, 1995.

49. Hansen OH, Pedersen T, and Larsen JK. Cell proliferation kinetics in normal human gastric mucosa. Studies on diurnal fluctuations and effect of food ingestion. *Gastroenterology* 70: 1051-1054, 1976.

50. Harley VR, Lovell-Badge R, and Goodfellow PN. Definition of a consensus DNA binding site for SRY. *Nucleic Acids Res* 22: 1500-1501, 1994.

51. Healy C, Uwanogho D, and Sharpe PT. Expression of the chicken Sox9 gene marks the onset of cartilage differentiation. *Ann N Y Acad Sci* 785: 261-262, 1996.

52. Helander HF. The cells of the gastric mucosa. *Int Rev Cytol* 70: 217-289, 1981.

53. Hosking BM, Muscat GE, Koopman PA, Dowhan DH, and Dunn TL. Trans-activation and DNA-binding properties of the transcription factor, Sox-18. *Nucleic Acids Res* 23: 2626-2628, 1995.

54. Hough MR, Rosten PM, Sexton TL, Kay R, and Humphries RK. Mapping of CD24 and homologous sequences to multiple chromosomal loci. *Genomics* 22: 154-161, 1994.

55. Hudson C, Clements D, Friday RV, Stott D, and Woodland HR. Xsox17alpha and -beta mediate endoderm formation in Xenopus. *Cell* 91: 397-405, 1997.

56. Iguchi H, Ikeda Y, Okamura M, Tanaka T, Urashima Y, Ohguchi H, Takayasu S, Kojima N, Iwasaki S, Ohashi R, Jiang S, Hasegawa G, Ioka RX, Magoori K, Sumi K, Maejima T, Uchida A, Naito M, Osborne TF, Yanagisawa M, Yamamoto TT, Kodama T, and Sakai J. SOX6 attenuates glucose-stimulated insulin secretion by repressing PDX1 transcriptional activity and is down-regulated in hyperinsulinemic obese mice. *J Biol Chem* 280: 37669-37680, 2005.

57. Iguchi H, Urashima Y, Inagaki Y, Ikeda Y, Okamura M, Tanaka T, Uchida A, Yamamoto TT, Kodama T, and Sakai J. SOX6 suppresses cyclin D1 promoter activity by interacting with beta-catenin and histone deacetylase 1, and its down-regulation induces pancreatic beta-cell proliferation. *J Biol Chem* 282: 19052-19061, 2007.

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

59. Kanai-Azuma M, Kanai Y, Gad JM, Tajima Y, Taya C, Kurohmaru M, Sanai Y, Yonekawa H, Yazaki K, Tam PP, and Hayashi Y. Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development* 129: 2367-2379, 2002.

60. Karam SM. Lineage commitment and maturation of epithelial cells in the gut. *Front Biosci* 4: D286-298, 1999.

61. Karam SM, and Leblond CP. Identifying and counting epithelial cell types in the "corpus" of the mouse stomach. *Anat Rec* 232: 231-246, 1992.

62. Kasimiotis H, Myers MA, Argentaro A, Mertin S, Fida S, Ferraro T, Olsson J, Rowley MJ, and Harley VR. Sex-determining region Y-related protein SOX13 is a diabetes autoantigen expressed in pancreatic islets. *Diabetes* 49: 555-561, 2000.

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

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

65. Kay R, Rosten PM, and Humphries RK. CD24, a signal transducer modulating B cell activation responses, is a very short peptide with a glycosyl phosphatidylinositol membrane anchor. *J Immunol* 147: 1412-1416, 1991.

66. Kay R, Takei F, and Humphries RK. Expression cloning of a cDNA encoding M1/69-J11d heat-stable antigens. *J Immunol* 145: 1952-1959, 1990.

67. Kedinger M, Duluc I, Fritsch C, Lorentz O, Plateroti M, and Freund JN. Intestinal epithelial-mesenchymal cell interactions. *Ann N Y Acad Sci* 859: 1-17, 1998.

68. Kedinger M, Lefebvre O, Duluc I, Freund JN, and Simon-Assmann P. Cellular and molecular partners involved in gut morphogenesis and differentiation. *Philos Trans R Soc Lond B Biol Sci* 353: 847-856, 1998.

69. Kedinger M, Simon-Assmann P, Bouziges F, and Haffen K. Epithelial-mesenchymal interactions in intestinal epithelial differentiation. *Scand J Gastroenterol Suppl* 151: 62-69, 1988.

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

71. Korinek V, Barker N, Moerer P, van Donselaar E, Huls G, Peters PJ, and Clevers H. Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nat Genet* 19: 379-383, 1998.

72. Kormish JD, Sinner D, and Zorn AM. Interactions between SOX factors and Wnt/beta-catenin signaling in development and disease. *Dev Dyn* 239: 56-68, 2010.

73. Kwok C, Weller PA, Guioli S, Foster JW, Mansour S, Zuffardi O, Punnett HH, Dominguez-Steglich MA, Brook JD, Young ID, and et al. Mutations in SOX9, the gene responsible for Campomelic dysplasia and autosomal sex reversal. *Am J Hum Genet* 57: 1028-1036, 1995.

74. Leblond CP, Greulich, R. C., & Pereira, J. P. M. Relationship of cell formation and cell migration in the renewal of stratified squamous epithelia. In: *Advances of Biology of Skin*1964, p. 36-67.

75. Lee ER, Trasler J, Dwivedi S, and Leblond CP. Division of the mouse gastric mucosa into zymogenic and mucous regions on the basis of gland features. *Am J Anat* 164: 187-207, 1982.

76. Lefebvre V, Dumitriu B, Penzo-Mendez A, Han Y, and Pallavi B. Control of cell fate and differentiation by Sry-related high-mobility-group box (Sox) transcription factors. *Int J Biochem Cell Biol* 39: 2195-2214, 2007.

77. Li XL, Eishi Y, Bai YQ, Sakai H, Akiyama Y, Tani M, Takizawa T, Koike M, and Yuasa Y. Expression of the SRY-related HMG box protein SOX2 in human gastric carcinoma. *Int J Oncol* 24: 257-263, 2004.

78. Liber D, Domaschenz R, Holmqvist PH, Mazzarella L, Georgiou A, Leleu M, Fisher AG, Labosky PA, and Dillon N. 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.

79. Lim SC, and Oh SH. The role of CD24 in various human epithelial neoplasias. *Pathol Res Pract* 201: 479-486, 2005.

80. Lioubinski O, Muller M, Wegner M, and Sander M. Expression of Sox transcription factors in the developing mouse pancreas. *Dev Dyn* 227: 402-408, 2003.

81. Liu YZ, Pei YF, Liu JF, Yang F, Guo Y, Zhang L, Liu XG, Yan H, Wang L, Zhang YP, Levy S, Recker RR, and Deng HW. Powerful bivariate genome-wide association analyses suggest the SOX6 gene influencing both obesity and osteoporosis phenotypes in males. *PLoS One* 4: e6827, 2009.

82. Malhi H, and Gores GJ. Cellular and molecular mechanisms of liver injury. *Gastroenterology* 134: 1641-1654, 2008.

83. Mall FP. A study of the structural unit of the liver. *American Journal of Anatomy* 5: 227-308, 1906.

84. Mathis D, Vence L, and Benoist C. beta-Cell death during progression to diabetes. *Nature* 414: 792-798, 2001.

85. Matsui T, Kanai-Azuma M, Hara K, Matoba S, Hiramatsu R, Kawakami H, Kurohmaru M, Koopman P, and Kanai Y. Redundant roles of Sox17 and Sox18 in postnatal angiogenesis in mice. *J Cell Sci* 119: 3513-3526, 2006.

86. Menge BA, Tannapfel A, Belyaev O, Drescher R, Muller C, Uhl W, Schmidt WE, and Meier JJ. Partial pancreatectomy in adult humans does not provoke beta-cell regeneration. *Diabetes* 57: 142-149, 2008.

87. Montgomery RK, Carlone DL, Richmond CA, Farilla L, Kranendonk ME, Henderson DE, Baffour-Awuah NY, Ambruzs DM, Fogli LK, Algra S, and Breault DT. Mouse telomerase reverse transcriptase (mTert) expression marks slowly cycling intestinal stem cells. *Proc Natl Acad Sci U S A* 108: 179-184, 2011.

88. Mori-Akiyama Y, Akiyama H, Rowitch DH, and de Crombrugghe B. Sox9 is required for determination of the chondrogenic cell lineage in the cranial neural crest. *Proc Natl Acad Sci U S A* 100: 9360-9365, 2003.

89. Mori-Akiyama Y, van den Born M, van Es JH, Hamilton SR, Adams HP, Zhang J, Clevers H, and de Crombrugghe B. SOX9 is required for the differentiation of paneth cells in the intestinal epithelium. *Gastroenterology* 133: 539-546, 2007.

90. Nir T, Melton DA, and Dor Y. Recovery from diabetes in mice by beta cell regeneration. *J Clin Invest* 117: 2553-2561, 2007.

91. Ordonez C, Screaton RA, Ilantzis C, and Stanners CP. Human carcinoembryonic antigen functions as a general inhibitor of anoikis. *Cancer Res* 60: 3419-3424, 2000.

92. Otsubo T, Akiyama Y, Yanagihara K, and Yuasa Y. SOX2 is frequently downregulated in gastric cancers and inhibits cell growth through cell-cycle arrest and apoptosis. *Br J Cancer* 98: 824-831, 2008.

93. Patel S, Rew DA, Taylor I, Potten CS, Owen C, and Roberts SA. Study of the proliferation in human gastric mucosa after in vivo bromodeoxyuridine labelling. *Gut* 34: 893-896, 1993.

94. Pevny LH, and Lovell-Badge R. Sox genes find their feet. *Curr Opin Genet Dev* 7: 338-344, 1997.

95. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29: e45, 2001.

96. Piper K, Ball SG, Keeling JW, Mansoor S, Wilson DI, and Hanley NA. Novel SOX9 expression during human pancreas development correlates to abnormalities in Campomelic dysplasia. *Mech Dev* 116: 223-226, 2002.

97. Pontiggia A, Rimini R, Harley VR, Goodfellow PN, Lovell-Badge R, and Bianchi ME. Sex-reversing mutations affect the architecture of SRY-DNA complexes. *EMBO J* 13: 6115-6124, 1994.

98. Potten CS, Booth C, Tudor GL, Booth D, Brady G, Hurley P, Ashton G, Clarke R, Sakakibara S, and Okano H. Identification of a putative intestinal stem cell and early lineage marker; musashi-1. *Differentiation* 71: 28-41, 2003.

99. Powell DW, Mifflin RC, Valentich JD, Crowe SE, Saada JI, and West AB. Myofibroblasts. II. Intestinal subepithelial myofibroblasts. *Am J Physiol* 277: C183-201, 1999.

100. Prasad NB, Biankin AV, Fukushima N, Maitra A, Dhara S, Elkahloun AG, Hruban RH, Goggins M, and Leach SD. Gene expression profiles in pancreatic intraepithelial neoplasia reflect the effects of Hedgehog signaling on pancreatic ductal epithelial cells. *Cancer Res* 65: 1619-1626, 2005.

101. Pruszak J, Ludwig W, Blak A, Alavian K, and Isacson O. CD15, CD24, and CD29 define a surface biomarker code for neural lineage differentiation of stem cells. *Stem Cells* 27: 2928-2940, 2009.

102. Que J, Okubo T, Goldenring JR, Nam KT, Kurotani R, Morrisey EE, Taranova O, Pevny LH, and Hogan BL. Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. *Development* 134: 2521-2531, 2007.

103. Riccio O, van Gijn ME, Bezdek AC, Pellegrinet L, van Es JH, Zimber-Strobl U, Strobl LJ, Honjo T, Clevers H, and Radtke F. 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.

104. Rietze RL, Valcanis H, Brooker GF, Thomas T, Voss AK, and Bartlett PF. Purification of a pluripotent neural stem cell from the adult mouse brain. *Nature* 412: 736-739, 2001.

105. Sagiv E, Memeo L, Karin A, Kazanov D, Jacob-Hirsch J, Mansukhani M, Rechavi G, Hibshoosh H, and Arber N. CD24 is a new oncogene, early at the multistep process of colorectal cancer carcinogenesis. *Gastroenterology* 131: 630-639, 2006.

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

107. Sammar M, Aigner S, and Altevogt P. Heat-stable antigen (mouse CD24) in the brain: dual but distinct interaction with P-selectin and L1. *Biochim Biophys Acta* 1337: 287-294, 1997.

108. Sanada Y, Yoshida K, Konishi K, Oeda M, Ohara M, and Tsutani Y. Expression of gastric mucin MUC5AC and gastric transcription factor SOX2 in ampulla of vater adenocarcinoma: comparison between expression patterns and histologic subtypes. *Oncol Rep* 15: 1157-1161, 2006.

109. Sangiorgi E, and Capecchi MR. Bmi1 is expressed in vivo in intestinal stem cells. *Nat Genet* 40: 915-920, 2008.

110. Sasaki T, Giltay R, Talts U, Timpl R, and Talts JF. Expression and distribution of laminin alpha1 and alpha2 chains in embryonic and adult mouse tissues: an immunochemical approach. *Exp Cell Res* 275: 185-199, 2002.

111. Sato T, van Es JH, Snippert HJ, Stange DE, Vries RG, van den Born M, Barker N, Shroyer NF, van de Wetering M, and Clevers H. Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* 469: 415-418, 2011.

112. Sato T, Vries RG, Snippert HJ, van de Wetering M, Barker N, Stange DE, van Es JH, Abo A, Kujala P, Peters PJ, and Clevers H. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* 459: 262-265, 2009.

113. Schilham MW, Oosterwegel MA, Moerer P, Ya J, de Boer PA, van de Wetering M, Verbeek S, Lamers WH, Kruisbeek AM, Cumano A, and Clevers H. Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. *Nature* 380: 711-714, 1996.

114. Seguin CA, Draper JS, Nagy A, and Rossant J. Establishment of endoderm progenitors by SOX transcription factor expression in human embryonic stem cells. *Cell Stem Cell* 3: 182-195, 2008.

115. Seymour PA, Freude KK, Dubois CL, Shih HP, Patel NA, and Sander M. A dosage-dependent requirement for Sox9 in pancreatic endocrine cell formation. *Dev Biol* 323: 19-30, 2008.

116. Seymour PA, Freude KK, Tran MN, Mayes EE, Jensen J, Kist R, Scherer G, and Sander M. SOX9 is required for maintenance of the pancreatic progenitor cell pool. *Proc Natl Acad Sci U S A* 104: 1865-1870, 2007.

117. Shaheen NJ. Advances in Barrett's esophagus and esophageal adenocarcinoma. *Gastroenterology* 128: 1554-1566, 2005.

118. Sinclair AH, Berta P, Palmer MS, Hawkins JR, Griffiths BL, Smith MJ, Foster JW, Frischauf AM, Lovell-Badge R, and Goodfellow PN. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* 346: 240-244, 1990.

119. Sinner D, Kordich JJ, Spence JR, Opoka R, Rankin S, Lin SC, Jonatan D, Zorn AM, and Wells JM. Sox17 and Sox4 differentially regulate beta-catenin/T-cell factor activity and proliferation of colon carcinoma cells. *Mol Cell Biol* 27: 7802-7815, 2007.

120. Smits P, Li P, Mandel J, Zhang Z, Deng JM, Behringer RR, de Crombrugghe B, and Lefebvre V. The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. *Dev Cell* 1: 277-290, 2001.

121. Snippert HJ, van Es JH, van den Born M, Begthel H, Stange DE, Barker N, and Clevers H. Prominin-1/CD133 marks stem cells and early progenitors in mouse small intestine. *Gastroenterology* 136: 2187-2194 e2181, 2009.

122. Soullier S, Jay P, Poulat F, Vanacker JM, Berta P, and Laudet V. Diversification pattern of the HMG and SOX family members during evolution. *J Mol Evol* 48: 517-527, 1999.

123. Spence JR, Lange AW, Lin SC, Kaestner KH, Lowy AM, Kim I, Whitsett JA, and Wells JM. Sox17 regulates organ lineage segregation of ventral foregut progenitor cells. *Dev Cell* 17: 62-74, 2009.

124. Spradling A, Drummond-Barbosa D, and Kai T. Stem cells find their niche. *Nature* 414: 98-104, 2001.

125. Stern CD, and Fraser SE. Tracing the lineage of tracing cell lineages. *Nat Cell Biol* 3: E216-218, 2001.

126. Stocker E, and Pfeifer U. [On the manner of proliferation of the liver parenchyma after partial hepatectomy. Autoradiography studies using 3H-thymidine]. *Naturwissenschaften* 52: 663, 1965.

127. Streit A, Sockanathan S, Perez L, Rex M, Scotting PJ, Sharpe PT, Lovell-Badge R, and Stern CD. Preventing the loss of competence for neural induction: HGF/SF, L5 and Sox-2. *Development* 124: 1191-1202, 1997.

128. Szutorisz H, and Dillon N. The epigenetic basis for embryonic stem cell pluripotency. *Bioessays* 27: 1286-1293, 2005.

129. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, and Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131: 861-872, 2007.

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

131. Tani Y, Akiyama Y, Fukamachi H, Yanagihara K, and Yuasa Y. Transcription factor SOX2 up-regulates stomach-specific pepsinogen A gene expression. *J Cancer Res Clin Oncol* 133: 263-269, 2007.

132. Taranova OV, Magness ST, Fagan BM, Wu Y, Surzenko N, Hutton SR, and Pevny LH. SOX2 is a dose-dependent regulator of retinal neural progenitor competence. *Genes Dev* 20: 1187-1202, 2006.

133. Thein DC, Thalhammer JM, Hartwig AC, Crenshaw EB, 3rd, Lefebvre V, Wegner M, and Sock E. The closely related transcription factors Sox4 and Sox11 function as survival factors during spinal cord development. *J Neurochem* 115: 131-141, 2010.

134. Thomas GA, Williams D, and Williams ED. The demonstration of tissue clonality by X-linked enzyme histochemistry. *J Pathol* 155: 101-108, 1988.

135. Tsukamoto T, Inada K, Tanaka H, Mizoshita T, Mihara M, Ushijima T, Yamamura Y, Nakamura S, and Tatematsu M. Down-regulation of a gastric transcription factor, Sox2, and ectopic expression of intestinal homeobox genes, Cdx1 and Cdx2: inverse correlation during progression from gastric/intestinal-mixed to complete intestinal metaplasia. *J Cancer Res Clin Oncol* 130: 135-145, 2004.

136. Uemura M, Hara K, Shitara H, Ishii R, Tsunekawa N, Miura Y, Kurohmaru M, Taya C, Yonekawa H, Kanai-Azuma M, and Kanai Y. Expression and function of mouse Sox17 gene in the specification of gallbladder/bile-duct progenitors during early foregut morphogenesis. *Biochem Biophys Res Commun* 391: 357-363, 2010.

137. Uwanogho D, Rex M, Cartwright EJ, Pearl G, Healy C, Scotting PJ, and Sharpe PT. Embryonic expression of the chicken Sox2, Sox3 and Sox11 genes suggests an interactive role in neuronal development. *Mech Dev* 49: 23-36, 1995.

138. van de Wetering M, and Clevers H. Sequence-specific interaction of the HMG box proteins TCF-1 and SRY occurs within the minor groove of a Watson-Crick double helix. *EMBO J* 11: 3039-3044, 1992.

139. van der Flier LG, Haegebarth A, Stange DE, van de Wetering M, and 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.

140. van der Flier LG, van Gijn ME, Hatzis P, Kujala P, Haegebarth A, Stange DE, Begthel H, van den Born M, Guryev V, Oving I, van Es JH, Barker N, Peters PJ, van de Wetering M, and Clevers H. Transcription factor achaete scute-like 2 controls intestinal stem cell fate. *Cell* 136: 903-912, 2009.

141. van Es JH, van Gijn ME, Riccio O, van den Born M, Vooijs M, Begthel H, Cozijnsen M, Robine S, Winton DJ, Radtke F, and Clevers H. Notch/gamma-secretase inhibition turns proliferative cells in intestinal crypts and adenomas into goblet cells. *Nature* 435: 959-963, 2005.

142. van Houte LP, Chuprina VP, van der Wetering M, Boelens R, Kaptein R, and Clevers H. Solution structure of the sequence-specific HMG box of the lymphocyte transcriptional activator Sox-4. *J Biol Chem* 270: 30516-30524, 1995.

143. Wagner T, Wirth J, Meyer J, Zabel B, Held M, Zimmer J, Pasantes J, Bricarelli FD, Keutel J, Hustert E, and et al. Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell* 79: 1111-1120, 1994.

144. Wang X, Foster M, Al-Dhalimy M, Lagasse E, Finegold M, and Grompe M. The origin and liver repopulating capacity of murine oval cells. *Proc Natl Acad Sci U S A* 100 Suppl 1: 11881-11888, 2003.

145. Wang X, Hu J, Zhao D, Wang G, Tan L, Du L, Yang J, Hou L, Zhang H, Yu Y, Deng H, and Ding M. NestinnegCD24low/- population from fetal Nestin-EGFP transgenic mice enriches the pancreatic endocrine progenitor cells. *Pancreas* 31: 385-391, 2005.

146. Wang Y, Giel-Moloney M, Rindi G, and Leiter AB. Enteroendocrine precursors differentiate independently of Wnt and form serotonin expressing adenomas in response to active beta-catenin. *Proc Natl Acad Sci U S A* 104: 11328-11333, 2007.

147. Weiss MA. Floppy SOX: mutual induced fit in hmg (high-mobility group) box-DNA recognition. *Mol Endocrinol* 15: 353-362, 2001.

148. Wells JM, and Melton DA. Vertebrate endoderm development. *Annu Rev Cell Dev Biol* 15: 393-410, 1999.

149. Wild S, Roglic G, Green A, Sicree R, and King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care* 27: 1047-1053, 2004.

150. Wilson ME, Yang KY, Kalousova A, Lau J, Kosaka Y, Lynn FC, Wang J, Mrejen C, Episkopou V, Clevers HC, and German MS. The HMG box transcription factor Sox4 contributes to the development of the endocrine pancreas. *Diabetes* 54: 3402-3409, 2005.

151. Wirth T, Soeth E, Czubayko F, and Juhl H. Inhibition of endogenous carcinoembryonic antigen (CEA) increases the apoptotic rate of colon cancer cells and inhibits metastatic tumor growth. *Clin Exp Metastasis* 19: 155-160, 2002.

152. Wood HB, and Episkopou V. Comparative expression of the mouse Sox1, Sox2 and Sox3 genes from pre-gastrulation to early somite stages. *Mech Dev* 86: 197-201, 1999.

153. Wright E, Hargrave MR, Christiansen J, Cooper L, Kun J, Evans T, Gangadharan U, Greenfield A, and Koopman P. The Sry-related gene Sox9 is expressed during chondrogenesis in mouse embryos. *Nat Genet* 9: 15-20, 1995.

154. Wright N, Allison, M. *The Biology of Epithelial Cell Populations*. Oxford: Clarindon, 1984.

155. Xu XL, Xing BC, Han HB, Zhao W, Hu MH, Xu ZL, Li JY, Xie Y, Gu J, Wang Y, and Zhang ZQ. The properties of tumor-initiating cells from a hepatocellular carcinoma patient's primary and recurrent tumor. *Carcinogenesis* 31: 167-174, 2010.

156. Yang Q, Bermingham NA, Finegold MJ, and Zoghbi HY. Requirement of Math1 for secretory cell lineage commitment in the mouse intestine. *Science* 294: 2155-2158, 2001.

157. Yuan H, Corbi N, Basilico C, and 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.

158. Zhou CJ, Guo JQ, Zhu KX, Zhang QH, Pan CR, Xu WH, Wang HJ, and Liu B. Elevated expression of SOX9 is related with the progression of gastric carcinoma. *Diagn Cytopathol* 2010.

159. Zou GM. Cancer initiating cells or cancer stem cells in the gastrointestinal tract and liver. *J Cell Physiol* 217: 598-604, 2008.