CALL FOR PAPERS | Intestinal Stem Cells in GI Physiology and Disease

Distinct levels of Sox9 expression mark colon epithelial stem cells that form colonoids in culture

Department of Medicine, Division of Gastroenterology and Hepatology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

Submitted 14 July 2011; accepted in final form 10 October 2011

Ramalingam S, Daughtridge GW, Johnston MJ, Gracz AD, Magness ST. Distinct levels of Sox9 expression mark colon epithelial stem cells that form colonoids in culture. Am J Physiol Gastrointest Liver Physiol 302: G10–G20, 2012. First published October 13, 2011; doi:10.1152/ajpgi.00277.2011.—Sox9 is an high-mobility group box transcription factor that is expressed in the stem cell zone of the small intestine and colon. We have previously used a Sox9EGFP mouse model to demonstrate that discrete levels of Sox9 expression mark small intestine epithelial stem cells that form crypt/villus-like structures in a three-dimensional culture system (Formeister EJ, Sionas AL, Lorance DK, Barkley CL, Lee GH, Magness ST. Am J Physiol Gastrointest Liver Physiol 296: G1108–G1118, 2009; Gracz AD, Ramalingam S, Magness ST. Am J Physiol Gastrointest Liver Physiol 298: G590–G600, 2010). In the present study, we hypothesized that discrete levels of Sox9 expression would also mark colonic epithelial stem cells (CESCs). Using the Sox9EGFP mouse model, we show that lower levels of Sox9 mark cells in the transit-amplifying progenitor cell zone, while higher levels of Sox9 mark cells in the colonic crypt base. Furthermore, we demonstrate that variable SOX9 levels persist in cells of colonic adenosas from mice and humans. Cells expressing lower Sox9 levels demonstrate gene expression profiles consistent with more differentiated populations, and cells expressing higher Sox9 levels are consistent with less differentiated populations. When placed in culture, cells expressing the highest levels of Sox9 formed “colonoids,” which are defined as bodies of cultured colonic epithelial cells that possess multiple crypt-like structures and a pseudolumen. Cells expressing the highest levels of Sox9 also demonstrate multipotency and self-renewal in vitro, indicating functional stemness. These data suggest a dose-dependent role for Sox9 in normal CESC and cells comprising colon tumors. Furthermore, distinct Sox9 levels represent a new biomarker to study CESC and progenitor biology in physiological and disease states.

SOX factors; colon epithelium culture; stem cell

For decades, the general region of the intestinal epithelial stem cell (IESC) niche in the small intestine and colon has been recognized to exist in the crypts (5, 31); however, the lack of specific biomarkers that can be used to identify and isolate IESCs has been a challenge to the study of IESCs during physiological states, disease, and injury. Recent studies have identified genes that exhibit restricted expression in the IESC zones of the small intestine, suggesting that they might serve as stem cell biomarkers (12, 26, 37, 39, 41). In vivo genetic lineage tracing studies in mice have demonstrated that the G protein-coupled receptor Lgr5, cell surface marker CD133, polycomb ring finger oncogene Bmi1, and telomerase (Tert) mark a population of IESCs in the crypt base of the small intestine. Additionally, the cell surface marker CD24 has been shown to be useful for fluorescent-activated cell sorting (FACS) enrichment of a small intestine IESC population that possesses stem cell characteristics, which are defined by multipotency and self-renewal capacity in vitro (15, 44). Other studies in the small intestine have correlated gene expression patterns of Axl2, Olfm4, Musashi1, and Dcamkl1 to IESC zone (17, 22, 30, 42, 43); however, it is still unclear whether these genes mark cells that possess functional stem cell properties based on in vivo lineage tracing or in vitro culture of FACS isolated cells.

Our laboratory has previously demonstrated that distinct Sox9 expression levels are useful biomarkers in the identification and isolation of different small intestine epithelial cell lineages, including IESCs (15). Using a Sox9EGFP mouse model in which enhanced green fluorescent protein (EGFP) is driven by large regulatory regions of the Sox9 5′-flanking regions, we demonstrated that “high” Sox9EGFP levels mark chromogranin A-positive cells, which are consistent with enteroendocrine lineages; “low” levels mark IESCs, and “sublow” levels mark transit-amplifying progenitor cells (10, 15). Furthermore, we demonstrated functional stem cell characteristics of Sox9EGFPlow cells by introducing these cells into a culture system that facilitates formation of crypt/villus-like structures from a single cell (15). Subsequently, genetic lineage tracing has shown Sox9 to mark multipotent IESCs in vivo (13).

While considerable efforts have been made to define the genetic signature of small intestine epithelial cells (2, 9, 10, 15, 17, 18, 25, 26, 30, 37, 41–44), fewer studies have focused on identifying genetic biomarkers of colonic epithelial stem cells (CESCs) (11, 13, 36, 39). The concept of two populations of CESC has been postulated: one is a slowly dividing label-retaining population, and the other is an actively dividing high-turnover population (19). Two genetic lineage-tracing studies describe Lgr5 and Sox9 as biomarkers of small IESCs and CESC (13, 39). Two other candidate biomarkers, β1-integrin and Phlda1 (Pleckstrin homology-like domain family A member 1), localize to the stem cell zone in the colon (11, 36). Phlda1 demonstrates a restricted expression pattern at the base of the crypt and appears to be consistent with Lgr5-expressing cells (36). FACS-isolated β1-integrin cells demonstrated the ability to form colonies in culture; however, at the time of this study, culture conditions did not permit testing of
whether the β1-integrin cells were multipotent and self-renewing (11). It is currently unknown whether any of these putative or established stem cell biomarkers demonstrate overlapping expression patterns with a label-retaining or high-turnover CESC population. Sox9 expression marks colonic epithelial stem cells. G11

EXPERIMENTAL PROCEDURES

whether the β1-integrin cells were multipotent and self-renewing (11). It is currently unknown whether any of these putative or established stem cell biomarkers demonstrate overlapping expression patterns with a label-retaining or high-turnover CESC population. Sox9 is regionally restricted to the progenitor and stem cell zones. Therefore, we hypothesized that distinct levels of Sox9 could serve as useful biomarkers for CESC and progenitor populations in the colon. This study describes the cellular and molecular characterization of CESC that express Sox9. We quantify different levels of Sox9 in single cells of the colonic crypt base and in cells of colon tumors from both mice and humans using confocal microscopy and computational image analysis. We assess the stemness of colonic epithelial cells exhibiting different Sox9 expression levels. Additionally, we define culture conditions that support colonoid formation from single CESC. These studies provide a strong framework for identification of other novel markers of CESC and facilitate investigation into the behavior of CESC in disease models of the colon.

EXPERIMENTAL PROCEDURES

Mice genotyping. The Sox9EGFP mouse line was originally generated as part of the GENSAT Brain Atlas Project (14) and contains genomic integration of a modified bacterial artificial chromosome (BAC) (RP32-149D18) with ~75.5-kb upstream and ~151-kb downstream sequence to Sox9. Frozen Sox9EGFP 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. Heterozygotes on the CD-1 genetic background. Mice breed normally and live to adulthood with no overt phenotypes due to the transgene. 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.

Epithelial cell dissociation FACS. The distal colons from three freshly killed 6- to 8-wk-old Sox9EGFP mice were removed, flushed of luminal contents with ice-cold Hank's balanced salt solution (HBSS), and incubated at 37°C for 1 h in 5 ml of colonic epithelium dissociation medium I (CEDM I; Dulbecco's modified eagle medium-high glucose (DMEM-H) containing 1 U/ml penicillin, 1 μg/ml streptomycin, 2.9 μg/ml l-glutamine, 1% fetal bovine serum (FBS), 5 mM calcium chloride, 150 U/ml collagenase XI (Sigma, St. Louis, MO), and 0.03 U/ml dispase (BD Biosciences, Franklin Lakes, NJ). The tissue was shaken vigorously for 1 min in CEDM III. The remnant tissue (muscularis) was then removed from CEDM III, and dissociated crypts were pelleted at 820 × g for 4 min. The crypts were washed in 10 ml ice-cold Dulbecco's phosphate-buffered saline (GIBCO), pelleted at 1,280 g, and incubated at 37°C for 10 min in CEDM IV (GIBCO). Crypts were shaken vigorously every 2 min to maximize single-cell sorting. Sort gates for CECS and facilitate investigation into the behavior of CECS in disease models of the colon.

Epithelial cell dissociation FACS. The distal colons from three freshly killed 6- to 8-wk-old Sox9EGFP mice were removed, flushed of luminal contents with ice-cold Hank's balanced salt solution (HBSS), and incubated at 37°C for 1 h in 5 ml of colonic epithelium dissociation medium I (CEDM I; Dulbecco's modified eagle medium-high glucose (DMEM-H) containing 1 U/ml penicillin, 1 μg/ml streptomycin, 2.9 μg/ml l-glutamine, 1% fetal bovine serum (FBS), 5 mM calcium chloride, 150 U/ml collagenase XI (Sigma, St. Louis, MO), and 0.03 U/ml dispase (BD Biosciences, Franklin Lakes, NJ). The tissue was shaken vigorously for 1 min in CEDM III. The remnant tissue (muscularis) was then removed from CEDM III, and dissociated crypts were pelleted at 820 × g for 4 min. The crypts were washed in 10 ml ice-cold Dulbecco's phosphate-buffered saline (GIBCO), pelleted at 1,280 g, and incubated at 37°C for 10 min in CEDM IV (GIBCO). Crypts were shaken vigorously every 2 min to maximize single-cell sorting. Sort gates for CECS and facilitate investigation into the behavior of CECS in disease models of the colon.

Abstract. Atoh1, Chga, and Lgr5 expression that is unique to the colonic epithelium, we normalized fold changes to population 1, which exhibited the lowest levels of Sox9EGFP expression. All data points represent means ± SE from three separate experiments; statistical analysis was conducted by one-way ANOVA and post hoc Tukey-Kramer pairwise tests. A P value of <0.05 was considered statistically significant.

CESC culturing intestinal myofibroblasts. E16.5 embryonic intestinal myofibroblasts (MFs) were isolated from a transgenic mouse model in which the red fluorescent reporter gene, dsRED is driven by the smooth muscle actin (SMA) promoter (21). A CD-1 female was time-mated with a SMA-DSRED male mouse. Embryos were collected at 16.5 days postcoitus and phenotyped for expression of SMA dsRed by visualization under fluorescent microscopy. Intestines were dissected from SMA dsRed-positive embryos with the aid of a dissecting microscope and rinsed in sterile, ice-cold PBS to remove any adherent nonintestinal tissues. Special care was taken to assure no embryonic liver remained in the cell prep. Tissue was minced into ~1-mm2 pieces and transferred to a 50-ml conical containing 5 ml DMEM-H (GIBCO) with 3.0 U/ml dispase (Collaborative Biomedical) and 300 U/ml collagenase (Sigma). The tissue slurry was rotated for 30 min at room temperature before 10 ml of DMEM-H + 10% FBS (Gemini) was added. The slurry was then pipetted vigorously 50 times to mechanically dissociate any remaining clumps of cells. The cells were centrifuged for 5 min at 460 g, resuspended in DMEM-H with 10% FBS, and plated in a 10-cm culture dish. Cells underwent three passages to obtain a pure population of dsRED-positive MFs before use in colonoid culture. Flow cytometry demonstrated that >98% of cells were dsRED-positive.

Ten thousand sorted cells from each gate (1, 2, or 3) were individually pelleted at 3,000 g for 5 min, the media was carefully removed, and 3 μl of media were added to the cell pellet, which was resuspended by pipetting. The cells were added to 50 μl of ice-cold growth factor reduced Matrigel (BD, Franklin Lakes, NJ) supplemented with 1 μg/ml of mR-Spondin1 (R&D Systems, Minneapolis, MN), 1 μM Jagged-1 (AnaSpec, San Jose, CA), 100 ng/ml Noggin...
(Peprotech, Rocky Hill, NJ), 80 mg/ml Wnt3a (R&D Systems, Minneapolis, MN), and 50 ng/ml epidermal growth factor (EGF) (R&D Systems) and thoroughly mixed by pipetting. The cell/Matrigel mixture was plated on a Transwell insert designed for a 24-well plate format. The Matrigel was allowed to polymerize for 30 min at 37°C. The Transwells were placed into the 24-well plate, which contained E18.5 MFs that were 30–50% confluent. Two-hundred and fifty microliters of CCM with 10 µM Y27632 were added to the well containing the MFs, and 200 µl of CCM were added to the Transwell insert (500-µl total volume). Growth factors were replenished every 2 days by diluting the factors (R-spondin1, Wnt3a, Noggin, EGF) in a volume of growth medium, such that 10 µl represented a 50× solution of concentrations described above. Growth medium containing R-spondin, Wnt3a, Noggin, and EGF at 1× concentrations was completely changed every 4 days.

For passaging experiments, colonoids were isolated from Matrigel with a pipette and transferred to advanced DMEM/F12 with 10 µM Y27632 in 2.0 ml conical tubes. Single cells were obtained by addition of dispase (0.3 U/ml), incubation for 20 min at 37°C, and combined with trituration 20× through a 10-ml pipette connected to a mechanical pipettor. If cells were not dissociated to satisfaction, another 20 triturations were performed. Cells were pelleted and washed twice in advanced DMEM/F12, containing 10% FBS and 10 µM Y27632. Cells were washed once in advanced DMEM/F12, pelleted at 3,000 g, and replated in Matrigel, as described above.

**Immunostaining/microscopy.** For tissue preparation, small intestines were dissected from adult Sox9EGFP mice (>8 wk of age), and luminal contents were flushed out with PBS, followed immediately by a single flush with freshly prepared 4% paraformaldehyde. The colon was opened along the proximal-distal axis and fixed for an additional 14–18 h at 4°C. The tissues were then prepared for cryosectioning by immersion in 30% sucrose solution for at least 24 h 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 (8–10 µm) 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. Human colon tumor samples were obtained from US Biomax (Rockville, MD). Twenty-four cases were contained in a tissue microarray (TMA; serial numbers T0482). Paraffin was removed from the TMA using Histoclear, and immunostaining was conducted as described for mouse tissues.

For immunostaining, the sections were washed twice in PBS to remove OCT, followed by incubation in blocking medium (5% normal goat or donkey serum), in PBS-0.3% Triton X-100 for at least 30 min at room temperature (21–25°C). Primary antibodies were applied to the tissue sections in antibody-staining solution (1% normal goat serum, in PBS-0.3% Triton X-100). Dilutions were as follows: α-lysozyme (rabbit; 1:1,000, Diagnostics BioSystems, Pleasanton, CA, RP 028), α-mucin 2 (rabbit; 1:100, Santa Cruz Biotechnology, Santa Cruz, CA, sc-15334), α-SOX9 (rabbit; 1:1,000, Chemicon, Temecula, CA, AB5535), α-substance P (rat; 1:500, Chemicon, MAB356), α-chromogranin A (rabbit; 1:500, Immunostar, 20086), and α-sucrase isomaltase (goat; 1:100, Santa Cruz Biotechnology, sc-27603). All secondary antibodies (α-rabbit-Cy3, Sigma, St. Louis, MO, C2306; α-rabbit-Alexa fluor 488, Molecular Probes, Eugene, OR, Z-25302; α-rat-Alexa fluor 555, Molecular Probes, Z-25305; α-goat-Cy3) were used at a 1:500 dilution in staining buffer. Ki67 immunostaining required special conditions. Tissue sections were cut as described above and were antigen retrieved using Reveel Buffer (Biocare Medical) for 10 s at 120°C in a pressure cooker. Slides were allowed to cool at room temperature for 20 min and washed 3× in PBS buffer (TB; 0.05 M Tris·HCl; pH 7.6) for 2 min each. Ki67 (DAKO) was diluted 1:100 in DAKO diluent (DAKO) and incubated in a humidified chamber for 2 h at room temperature. The Ki67 antibody solution was removed, and the tissues were washed 3× in TB for 2 min each. Biotinylated anti-rat-IgG was diluted 1:100 in DAKO diluent and applied to the tissue sections for 1 h at room temperature in a humidified chamber. The secondary antibody was removed, and the tissue sections were washed 3× in TB, as described above. For Ki67 detection, Alexa-fluor 555 streptavidin was diluted 1:250 in DAKO diluent and incubated with tissues for 1 h at room temperature in a humidified chamber. The slides were then washed 3×, as described above, and mounted using aqueous mounting medium. All nuclei were stained with Draq5 (1:10,000, Biostatus, San Diego, CA, BOS-889–001). Background staining was negligible, as determined by nonspecific IgG staining.

Day-to-day development of colonoids was tracked using light microscopy. Epifluorescence images were captured on an Olympus IX70 fitted with an Olympus digital camera. Objective lenses used were ×20 and ×40, with numerical apertures of 0.55 and 1.40, respectively. All confocal images represent ×200 (1.0-µm optical sections), ×400 (1.0-µm optical sections), or ×630 (1.0-µm optical sections). Objective lenses for the confocal images had a numerical aperture of 0.1.

**Image analysis.** Images for quantifying individual cell fluorescence intensities were obtained using a Zeiss LSM 510 confocal microscope at ×400 optical zoom with a 1.0-µm optical section. The red, green, and blue images were stacked in the image analysis program, FIJI (NIH Image). To measure the level of SOX9 immunostaining in individual cells, an oval was superimposed on the nucleus (blue channel), and the “measure” feature was used to sequentially determine the average fluorescence intensity (or pixel intensity) of the red and green wavelengths. The base of the crypt was defined as cell position 1, increasing to the right and decreasing to the left. Measurement data were then transferred into Microsoft Excel for statistical analysis. For each image, average fluorescence intensities were loaded into the analysis application “R” (R Foundation for Statistical Computing, Vienna, Austria), and the kernel density estimate was computed and plotted using the “kernel” function with Gaussian kernel and bandwidth set to 4.

**RESULTS**

Sox9 expression levels vary between cells in the normal colonic epithelium. Sox9 is expressed at different levels in the small intestine epithelium, facilitating the histological identification and FACS isolation of IESCs (10, 15). In the colon, Sox9EGFP is primarily expressed in the lower one-half of the crypt and is regionally separated from the more differentiated absorptive colonocytes, which express carboxic anhydrase II (Fig. 1A). To determine whether Sox9 was expressed at different levels in the colonic epithelium, we used a Sox9EGFP reporter gene mouse or immunostaining for endogenous SOX9 to visualize Sox9 expression (Fig. 1B). The data qualitatively suggest that Sox9EGFP (and SOX9) are expressed at variable levels, with higher Sox9EGFP expression in the crypt base and lower Sox9EGFP expression in cells positioned closer to the luminal surface (Fig. 1A).

To quantitatively assess whether cells in the crypt base expressed higher levels of Sox9EGFP and SOX9, image analysis software was used to analyze 1.0-µm confocal optical sections of Sox9EGFP colon that had been stained for endogenous SOX9 (Fig. 2). The nuclei of individual cells in the images were digitally marked (independent of Sox9 expression), and then Sox9EGFP and endogenous SOX9 levels in the marked nuclear regions of corresponding images were quantified. Comparisons were conducted on a cell-by-cell basis using the conventional crypt cell position method, where position “zero” is at the base of the crypt, cells to the left of zero are sequentially designated by negative numbers, and cells to the
right of zero are sequentially designated as positive numbers. Single crypt analysis shows that endogenous SOX9 levels trend with Sox9EGFP levels both at the single cell level and by location (Fig. 2A). Analyses of 20 crypts show that higher levels of Sox9EGFP and SOX9 are expressed toward the base of the crypt, and lower levels are expressed toward the luminal surface (Fig. 2B). The nonparametric Wilcoxon rank sum test supports the hypothesis that the distribution of Sox9EGFP intensities is significantly higher in the crypt base from cell positions $-4$ to $+4$. Together, these data indicate that Sox9EGFP/SOX9 expression in colonic crypt cells is indeed variable, with higher levels in the crypt base cells (from positions $-4$ to $+4$) and lower levels in cells toward the lumen (positions higher than $-4/4$), suggesting that different Sox9 levels could serve as useful biomarkers for distinct and localized cell lineages.

Colonic tumor cells exhibit variable SOX9- or Sox9EGFP-expression levels. SOX9 is expressed in colonic adenomas in mice and humans (3, 4, 20, 27, 33); however, Sox9 expression levels in individual cells have not been examined. To determine whether differential expression of Sox9EGFP persisted in murine adenomas, we treated Sox9EGFP mice with azoxymethane (AOM)/dextran sodium sulfate (DSS) and examined the levels of Sox9EGFP in developed tumors. Confocal images of adenomas visually show positive Sox9EGFP expression, while closer examination demonstrates differential Sox9EGFP expression levels in cells of the tumor (Fig. 3A). Statistical kernel density estimation of the distribution of Sox9EGFP is multimodal, indicating multiple levels of expression exist (Fig. 3B). To determine whether differential levels of SOX9 were expressed in primary human tumors, we assessed SOX9 expression levels in individual cells from TMAs containing primary tumor biopsies (Fig. 3C). Again, statistical kernel density estimation of the distribution of SOX9 expression is multimodal, indicating all three human tumors show multiple, differential levels of expression (Fig. 3D).

Sox9 is expressed in undifferentiated proliferating cells and a subset of mucin 2 expressing cells. Like most SOX9-expressing cells in the small intestine crypts, a majority of the SOX9-expressing cells in colonic crypts also express the Ki67 proliferation marker (Fig. 4A). We identified a minority subset of SOX9-expressing cells that did not colocalize with the Ki67 proliferation marker (Fig. 4A; white arrows). The basal nuclei and large mucous globules in the cytoplasm suggest these nonproliferating cells are goblet cells (6). To determine the lineage identity of Sox9EGFP-expressing cells, we stained Sox9EGFP colonic tissue for goblet and enteroendocrine cell biomarkers. Immunostaining demonstrates that enteroendocrine cells (chromogranin A$^+$; Fig. 4B) and goblet cells (MUC2$^+$; Fig. 4C) do not express detectable levels of
SOX9 EXPRESSION MARKS COLONIC EPITHELIAL STEM CELLS

Fig. 2. Confocal and computational image analysis demonstrates higher Sox9EGFP/Sox9 expression levels in the colonic crypt base. A: single-cell image analysis. Sox9EGFP and Sox9 immunofluorescence intensities were measured on a cell-by-cell basis and plotted as a function of cell position. Spearman rank correlation test indicated Sox9EGFP intensities correlated with endogenous Sox9 fluorescence intensities by a significant dependence between endogenous Sox9 and enhanced green fluorescent protein (EGFP) intensity distributions ($r_s = 0.69$ with $P < 2.2 	imes 10^{-16}$). B: Sox9EGFP levels form distinct subgroups. Fluorescence intensities were averaged across 20 crypts for each cell position. When organized by location into two subgroups at the crypt base (positions $-4$ to $+4$; solid circle) and higher (positions $+5$ to $+10$; shaded circle), the distribution of Sox9EGFP intensities was significantly higher in the crypt base group compared with the TA (transit amplifying) zone group (Wilcoxon rank sum test, $P = 1.214 	imes 10^{-9}$ for EGFP, $P = 3.482 	imes 10^{-9}$ for Sox9).

Sox9EGFP. Although the absence of endogenous Sox9 expression is consistent with the absence of Sox9EGFP transgene in enteroendocrine cells (data not shown), a subset of MUC2+/Sox9EGFP-goblet cells do express endogenous Sox9 (Fig. 4D). Thus, in the colonic epithelium, Sox9 is expressed in CESC/progenitors and a subset of goblet cells. Despite the inconsistent expression pattern of the Sox9EGFP transgene and endogenous Sox9 in the crypt-based MUC2+ populations, the transgene faithfully recapitulates the endogenous gene in the proliferating stem/progenitor cell populations. The lack of Sox9EGFP expression in the crypt-based MUC2+ cells was exploited as a method to FACS exclude MUC2+ cells from putative CESC populations.

Increasing expression levels of Sox9 in colonic epithelial cells correlate with increased stemness. In the small intestine epithelium, Sox9EGFP is expressed at different levels in the base of the crypts, and distinct levels of EGFP fluorescence facilitate differential identification and FACS isolation of IESCs from transit amplifying progenitors and enteroendocrine cells (15). To assess whether different levels of Sox9EGFP expression observed in the colonic epithelium could be used in a similar manner to isolate CESC, flow cytometry was used to analyze the Sox9EGFP levels in dissociated colonic epithelial cells (Fig. 5). The results show three different Sox9EGFP-expressing populations, which are defined by three distinct EGFP peak intensities (Fig. 5A). “Population 1” was defined by cells expressing the lowest levels of Sox9EGFP, “population 2” by intermediate Sox9EGFP levels, and “population 3” by cells expressing the highest levels of Sox9EGFP.

To characterize Sox9EGFP-expressing populations at the transcript level, FACS was used to differentially isolate the three Sox9EGFP-expressing populations, and lineage identity was assessed by real-time RT-PCR analysis of each population. Increasing endogenous Sox9 mRNA expression correlated with increasing Sox9EGFP expression validating the FACS (Fig. 5, B and C). Previous studies using genetic lineage tracing have identified that Lgr5 marks a CESC population in vivo (39). Population 3, which is most enriched for Sox9 mRNA, also demonstrates significant enrichment of Lgr5 mRNA, compared with population 1, suggesting that higher levels of Sox9EGFP correlate with increased “stemness” in the colon (Fig. 5D). Notch pathway activation has been demonstrated to be essential for maintaining stemness (16, 35, 40). When the Notch pathway is downregulated, there is an increase in Atoh1 expression, which results in secretory lineage specification in transit amplifying progenitor cells. The data demonstrate a modest but significant decrease in Notch1 mRNA in the population 1, which expresses the lowest level of Sox9, and more notably a significant increase in secretory progenitor cell marker, Atoh1, mRNA in population 1. These data indicate that lower levels of Sox9EGFP mark cells consistent with transit amplifying progenitor cells (Fig. 5, E and F). Additionally, there was increased mRNA expression associated with goblet cells (Tff3) and enteroendocrine cells (chromogranin A) in cells that expressed the lowest levels of Sox9EGFP (Fig. 5, G and H). It is noteworthy that neither goblet cells nor enteroendocrine cells are observed to express Sox9EGFP by fluorescence microscopy; however, the high sensitivity of EGFP detection by flow cytometry is likely detecting small amounts of EGFP, even in cells that have downregulated Sox9EGFP and upregulated Muc2 and Chga upon differentiation. Taken together, these data strongly suggest that higher levels of Sox9EGFP/Sox9 mark CESC, and lower levels mark differentiating progenitors.

Colonic epithelial cells expressing the highest level of Sox9EGFP by flow cytometry form colonoids in culture. Recent advances in culturing techniques for small intestine IESCs have revolutionized single-cell analysis of putative IESC populations that are initially identified and FACS isolated based on their gene expression signatures. It should be stated that a true clonogenic assay in which validated single cells are FACSed into individual wells has not been developed for IESCs, but, rather, single cells are typically FACS isolated and plated together as putative singlets into a three-dimensional culture system. Sato et al. (39) first described the culture system that supported the development of crypt/villus-like structures from single IESCs in the absence of mesenchymal support cells. These culture conditions provide essential signaling that is usually supplied by the nonepithelial cells in vivo, namely, a laminin-enriched extracellular matrix, EGFP: Wnt-pathway agonist, R-Spondin1; Notch-receptor agonist, Jagged; and bone morphogenetic protein antagonist, Noggin (39). A subsequent study from the same group indicated that a “support cell” is...
likely required to produce mitogens and morphogens to FACS-
 isolated IESCs, which would promote survival and growth of
 single IESCs in culture. In the absence of a more precise
 clonogenic assay, we utilized the strategy developed by Sato et
 al. (39) to assess the stemness of Sox9EGFP-expressing
 colonocytes.

To test whether distinct Sox9 levels correlated with func-
tional stemness, we FACS isolated single Sox9EGFP cells
from populations 1, 2, and 3 and introduced these cells into the
culture conditions that support the development of small intes-
tine organoids. Forty-eight hours postplating, no viable cells
remained, based on absence of Sox9EGFP expression and
morphological characteristics consistent with widespread apop-
tosis. Recent studies indicate that a Paneth cell-derived
WNT3a source is critical for the high-efficiency development
of small intestine organoids from single IESCs (38). Paneth

Fig. 3. Differential Sox9EGFP/SOX9 expression levels persist in cells of azoxymethane (AOM)/dextran sodium sulfate (DSS) tumors in mice and in cells of
primary human colon tumors. A: variable Sox9EGFP expression levels are preserved in AOM/DSS adenomas. Region depicted in white box is magnified in right
panel. Images represent 1-μm optical sections imaged at ×200 (left) and ×400 (right) original magnification. B: Sox9EGFP fluorescence was quantified at the
single-cell level using the image analysis strategy described in Fig. 2. Multiple peaks in the histogram indicate a multimodal distribution of different Sox9EGFP
expression levels. C: three primary colon tumors from separate individuals were immunostained for SOX9. The data visually demonstrate variable levels of SOX9
in tumor cells. D: single-cell quantification of SOX9 expression in colon tumors demonstrates multimodal SOX9 expression and also different ratios of lower
and higher SOX9-expressing cells between tumors from different individuals.
cells are not present in the colonic epithelium; however, we reasoned that a “Paneth cell equivalent” might be providing WNT3a in vivo. The conditions that support generation of organoids from IESCs were modified to include exogenously supplied recombinant WNT3a, as well as an underlying layer of MFs derived from embryonic murine intestine. Pericryptal MFs have been implicated in the support of epithelial growth and development by secreting mitogens and morphogens in the IESC zones (23, 32). These new conditions supported development of mature colonoids, which formed from isolated whole crypts (Fig. 6A) and FACS-isolated Sox9EGFP cells from population 3 (Fig. 6B). Early development of colonospheres, a colonoid precursor defined as a multicellular sphere (with no crypt-like structures) possessing a pseudolumen, developed from populations 1, 2, and 3 at an incidence of 0.3, 0.6, and 1.1%, respectively, between 2 and 3 days postplating. However, by day 7, all colonospheres from populations 1 and 2 ceased further development and died, while all colonospheres from population 3 progressed development to colonoid stage by day 12. Colonoid development persisted through day 24, at which point the colonoids were mechanically dissociated and replated. Each colonoid at day 12 derived from FACS-isolated CESCs produced three independent colonoids upon passaging, indicating symmetric division and self-renewal capacity of population 3 CESCs (data not shown).

To assess the multipotency of population 3 CESCs, colonoids were sectioned and immunostained for markers of postmitotic lineages. Sox9EGFP expression persisted in colonoids and showed variable expression levels similar to those observed in vivo (Fig. 7). Epithelial cell adhesion molecule staining demonstrated that all cells were epithelial in nature (Fig. 7A). Chromogranin A and MUC2 expression indicated that population 3 CESCs were competent to generate enteroendocrine (CHGA+ and goblet cells (MUC2+) (Fig. 7, B and C). MUC2 staining in the pseudolumen of the colonoids also indicates that the cells derived from single CESCs are polarized and functional (Fig. 7C; middle). Carbonyl hydrazide II expression is consistent with absorptive colonocytes that localize to more luminal regions in vivo (Fig. 7D; see Fig. 1A). Together these data indicate that population 3 CESCs characterized by the highest Sox9 levels are multipotent and self-renewing.

**DISCUSSION**

Despite recent and rapid advances in the identification of stem cell biomarkers across nearly all tissue types, precise cell-specific gene expression levels and the phenotypic consequences of these levels remain an understudied metric of mechanisms regulating stem cell maintenance and differentiation. Technologies that enable detection and monitoring of a gene’s expression are becoming more robust, and the threshold of these detection methods is sufficient to enable quantification of gene expression in single cells. We propose that precise gene expression levels, and not simply the “on/off” qualitative parameter of gene expression, may be a more robust and critical identifier of cellular phenotype than previously appreciated.

In past studies, we have demonstrated this concept by using differential Sox9 expression at the cellular level to identify and isolate stem and progenitor cells from the small intestine epithelium (15). Other independent studies have used differential levels of Lgr5EGFP expression in the small intestine to demonstrate that “high” Lgr5EGFP-expressing cells are more stem-like, and “low” Lgr5EGFP-expressing cells are more consistent with transit amplifying progenitor cells (43). In the present study, we further validate this concept by showing that higher Sox9-expression levels mark cells that have gene expression signatures and functional characteristics of CESCs, while lower Sox9-expression levels are consistent with more differentiated cell types. Importantly, we also demonstrate that the presence of Sox9 expression in colonic epithelial cells is not, in and of itself, a sufficient indicator of stemness, as...
progenitor cells and subsets of goblet cells also express Sox9. These findings indicate that, while in vivo lineage tracing remains a powerful assay of stemness, careful analysis of cell-specific gene expression levels and the phenotypes associated with these levels are necessary criteria when defining and validating stem cell biomarkers.

We have demonstrated that colonic epithelial cells expressing the highest levels of Sox9EGFP also coexpress the highest levels of Lgr5, suggesting that these cells have functional stem cell qualities. Recent breakthroughs in culturing of isolated IESCs from the small intestine provided us with a framework to identify conditions that would support colonoid growth (15, 38, 39). The conditions described by Sato et al. (38, 39) for culture of small intestine IESCs did not support growth of colonoids from CESCs. We identified that additional requirements of WNT3a, in combination with critical unidentified diffusible factors from an underlying MF layer, were essential to support generation of colonoids from single cells. WNT3a is a secreted signaling protein that binds to receptors of the Frizzled family and has critical developmental roles in proliferation control and morphogenesis (24). Additionally, WNT3a has been shown to promote tissue-specific stem cell proliferation and maintenance in vitro (28, 34, 45). A recent study suggests that Paneth cells in the small intestine supply WNT3a to the small intestine IESCs and are required for maintenance of the IESC niche (38). Our studies demonstrate that WNT3a is essential for the growth of CESCs in culture, and we speculate that the WNT3a source in the colon could be generated from a similar cell type that is intimately associated with the CESCs. Our data suggest that this cell type is consistent with a Sox9-expressing goblet cell that is intimately connected to the CESCs in a manner similar to that of Paneth cells and IESCs in the small intestine. Moreover, both the crypt-based goblet cells and Paneth cells express Sox9, indicating a conserved role for this transcription factor in differentiated cell types that comprise the stem cell zone. Interestingly, the Sox9EGFP transgene is not expressed in either MUC2+ cells in the base of the colonic crypt or Paneth cells in the small intestine, despite the expression of endogenous SOX9 in these cells types (10, 15). The Sox9EGFP transgene is randomly integrated into the genome; thus the disparity between the Sox9EGFP transgene expression and endogenous SOX9 expression could be due to genome position effects. It is also possible that the regulatory elements for Sox9 expression in crypt-based Paneth cells and goblet cell types are not contained in the BAC construct. If the latter is true, then...
efforts to identify the cis-regulatory elements required for lineage-specific Sox9 expression in Paneth cells and crypt-based goblet cell expression could be focused to genomic regions that flank the BAC ends. Identifying these regulatory elements could provide important insights into upstream signaling mechanisms that maintain Sox9 in cell types that support the IESC and CESC niche.

The ability to identify and isolate cancer cells based on their gene expression profiles and associated phenotypic characteristics has been central to the study of carcinogenesis (7, 46). Currently, the nature of how individual cells contribute to colon cancer growth and invasion remains controversial, as technological limitations have precluded the ability to precisely identify colon cancer stem cells. In this study and others, we have shown that Sox9EGFP levels vary between cell types with different lineage identities and functional stem cell characteristics in normal physiology (10, 15). A novel finding from our studies is that these variable levels are preserved in cells of AOM/DSS mouse colon tumors and primary human tumors. Based on the result from this study that higher levels of Sox9 expression correlate with increased stemness in physiological states, the quantitative assessment of the ratio of higher vs. lower Sox9-expressing cells within a tumor may provide an indication of cancer cell stemness. Further investigation into the levels of SOX9 in cancer cell stemness holds promise as a diagnostic tool to predict outcomes and direct therapeutic interventions.

Prevailing evidence suggests that Sox9 has anti-oncogenic properties, in that higher levels suppress proliferation and decrease Wnt signaling activity (4, 10). Interestingly, a new splice variant of SOX9 has been recently identified in human cancer cell lines and primary tumors that appear to exhibit oncogenic properties (1). The SOX9 splice variant (termed miniSox9) retains intron 2, producing in a premature stop codon and an aberrant COOH-terminus. The mutant protein retains the DNA binding domain, but lacks the transactivation domain. MiniSox9 inhibits the anti-oncogenic properties of full-length SOX9 and stimulates the canonical Wnt pathway, presumably increasing proliferative capacity and tumorigenicity (1). The miniSox9 study highlights the importance of characterizing the levels of Sox9 at the cellular level and also

Fig. 6. High Sox9EGFP expression can be used as a FACS parameter to isolate CESC that form colonoids in culture. Colonoids are defined as complex crypt-like bodies that develop from whole crypts or single CESC. Colonoids are composed of colonic epithelial cells surrounding a pseudolumen and do not contain a mesenchymal cellular component. By 3 days in culture, both whole colonic crypts (A) and single CESC (B) first develop into colonospheres that are composed of epithelial cells that have not formed crypt-like structures but contain a pseudolumen. Colonospheres progress to colonoids by 24 days in culture. Images were captured at ×200 original magnification.

Fig. 7. Sox9EGFP CESC that form colonoids are multipotent. A: epithelial cell adhesion molecule (EpCAM) staining confirms that colonoid cells are epithelial. ChgA (white arrow; B); MUC2 (white box indicates magnified region; C, right); and CAII (D) immunostaining indicate Sox9EGFP CESC are competent to generate enteroendocrine, goblet, and absorptive colonocytes, respectively. Images represent 1-μm optical sections. Magnifications are indicated in each panel.
understanding how these Sox9 levels mechanistically confer cellular phenotype. For example, determining expression levels of full-length or miniSox9 in the same cell may identify subpopulations of cells within the colonic tumors that have differential capacity for proliferation and differentiation. The Sox9EGFP mouse model will be a useful tool to study the role of differential Sox9-expression levels in cells of colon tumors. Currently, the Sox9EGFP mice are on a genetic background that is highly resistant to AOM/DSS tumor induction; thus the amount of tumor tissue obtained in this study was insufficient for gene expression analysis and culture assays at the single-cell level. On-going studies are underway to introduce the Sox9EGFP allele onto genetic backgrounds that are more amenable to colonic tumor induction. Based on our studies, we speculate that colonic adenoma cells expressing higher levels of Sox9 may exhibit slower division and more stem cell-like characteristics that have been hypothesized to be more resistant to chemotherapy and radiation, while lower levels of Sox9 would be consistent with more rapidly dividing progenitor-like cells that may be more responsive to standard cancer therapies (8). The concepts and image analysis strategies developed in this study will enable detailed analysis of Sox9 levels in human colonic adenomas and carcinomas and will likely establish rationale for screening tumor TMAs for different levels of oncogenic gene products and correlating cellular expression levels of these oncogenic proteins to clinical outcomes.

ACKNOWLEDGMENTS

We thank Ian Williamson for technical assistance, and Drs. Jim Simmons and P. Kay Lund for providing the protocol that we adapted for embryonic MF isolation and propagation. Additionally, we thank Dr. Susan Henning, Dr. Victoria Newton, Dr. Laurianne Van Landeghem, and Brent Puthoff for critical reading of the manuscript, and Dr. P. Kay Lund for useful discussions on terminology.

GRANTS

We acknowledge the University of North Carolina (UNC) Center for GI Biology and Disease Imaging and Histology COREs (P30-DK034987), the UNC Neuroscience Confocal Imaging Facility (P30-NS045892), the UNC Chapel Hill Mouse Regional Resource Center, and the UNC Flow Cytometry CORE. Funding was provided by the National Institutes of Health, K01- DK080181; the American Gastroenterological Association Research Scholar Award; the North Carolina Biotechnology Center; the UNC-Chapel Hill Center for Gastrointestinal Biology and Disease, P30-DK034987; and the UNC SPORE in GI Cancer, 5-P50-CA 106991.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS


REFERENCES


