Sox4 drives intestinal secretory differentiation toward tuft and enteroendocrine fates

Adam D Gracz[1], Matthew J Fordham[2], Danny C Trotier[2], Bailey Zwarycz[3], Yuan-Hung Lo[4], Katherine Bao[5], Joshua Starmer[1], Noah F Shroyer[4], Richard L Reinhardt[5]†, Scott T Magness[2,3,6*]

Affiliations:
[1] Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC
[2] Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC
[3] Department of Cell Biology & Physiology, University of North Carolina at Chapel Hill, Chapel Hill, NC
[4] Department of Medicine, Section of Gastroenterology and Hepatology, Baylor College of Medicine, Houston, TX
[5] Department of Immunology, Duke University, Durham, NC
[6] Joint Department of Biomedical Engineering, University of North Carolina at Chapel Hill/North Carolina State University, Chapel Hill, NC

*Correspondence to Lead Contact, Scott T. Magness: magness@med.unc.edu

† Current address: Department of Biomedical Research National Jewish Health and Department of Immunology and Microbiology, University of Colorado School of Medicine, Denver CO.

Keywords: intestinal stem cells, tuft cells, differentiation, Sox factors, Notch signaling

Abstract: Intestinal stem cells (ISCs) undergo differentiation following reduction of Notch signaling and activation of *Atoh1,* a master regulator of secretory fate. However, recent evidence suggests that rare tuft cells may differentiate via alternative mechanisms. Here, we examine the role of *Sox4* in ISC fate decisions. Loss of *Sox4* results in disruption of ISC function and secretory differentiation, including decreased numbers of enteroendocrine and tuft cells. *Sox4* is activated following reduction of Notch signaling and required for response to signals that induce tuft cell hyperplasia, including IL-13 and parasitic helminth infection. Subsets of *Sox4*+ progenitors are transcriptomically similar to both actively dividing and mature tuft cells, and dissimilar from other secretory progenitors expressing *Atoh1*. Our data suggest that *Sox4* drives an alternative, Notch-repressed secretory differentiation pathway independently of *Atoh1*, and challenge the status of *Atoh1* as the sole initiator of secretory fate in the intestine.

Main Text:

Introduction

The intestinal epithelium is essential for both digestive and barrier function and exhibits a multitude of physiological functions, including absorption of nutrients, secretion of hormones and antimicrobial peptides, response to infectious agents, and regeneration following physical damage. These functions are carried out by a diverse complement of post-mitotic cells, which can be broadly subdivided into: (a) absorptive enterocytes and (b) secretory lineages, which include antimicrobial-producing Paneth cells, mucus-producing goblet cells, and hormone-secreting enteroendocrine cells (EECs) (1). Tuft cells, which represent a less well-characterized lineage, initiate immune responses to parasitic infections in the intestine (2-5).With the exception of long-lived Paneth cells, these post-mitotic lineages are subject to the rapid (5-7 day) and continuous turnover of the intestinal epithelium. Therefore, their numbers must be maintained through constant proliferation and differentiation of the intestinal stem cells (ISCs), which reside in a specialized stem cell niche at the base of the intestinal crypts.

Extrinsic signals from the niche induce intrinsic regulatory programs in ISCs and their immediate progeny, transit-amplifying progenitor cells (TAs), to drive commitment to different cellular lineages. The Notch signaling pathway, which relies on cell-cell contact for forward and reverse signaling, is known to promote ISC and absorptive cell fates (6). Loss of Notch signaling in ISCs and early progenitors is associated with the induction of secretory differentiation through the derepression of *Atoh1*, which is widely regarded as the master regulator of secretory fate, driving formation of all Paneth cells, goblet cells, and EECs (6, 7). While tuft cells are also derived from *Lgr5*+ ISCs, the

transcriptional regulatory mechanisms governing their differentiation and hyperplastic response to infection remain less well characterized (3). This is due in part to evidence that tuft cells differentiate independently of *Atoh1*, but appear to require downregulation of Notch, presenting an apparently paradoxical Notch-repressed, *Atoh1*-independepent mechanism for tuft cell specification (7-9).

Sry-box containing (Sox) factors are transcription factors with broad regulatory roles in stem cell maintenance and differentiation in many adult tissues, including the intestinal epithelium (10, 11). *Sox9* is known to contribute to Paneth cell differentiation, and previous work from our group has demonstrated that distinct expression levels of *Sox9* can be used to differentially identify and isolate stem/progenitor cell populations from intestinal crypts (12-15). We extended this work to demonstrate that the highest levels of *Sox9* are associated with label-retaining cells (LRCs) and that *Sox9* is required to maintain ISC function following sub-lethal irradiation injury (16). Though Sox factors often exhibit complementary/compensatory function and dose-dependent activity(11), the role of other Sox factors in the adult intestinal epithelium remains undefined (11).

Sox4 has been associated with the ISC genomic signature and is correlated with increased invasion/metastasis and decreased survival in colon cancer, but its role in intestinal epithelial homeostasis is unknown (17, 18). In the hematopoietic system, as well as the developing nervous system, pancreas, kidney, and heart, *Sox4* is required for both proper cell lineage allocation and maintenance of progenitor pools (19-21). The diverse regulatory potential of *Sox4*, coupled with its known role in a broad range of tissue-specific progenitor populations, led us to hypothesize that it may function as a regulator of ISC maintenance and differentiation. In this study we establish a role for

Sox4 in ISC homeostasis and regulation of tuft cell hyperplasia during parasitic helminth infection, which represents a novel pathway for tuft cell specification.

Results

SOX4 is expressed at distinct levels in the ISC zone

To characterize *Sox4* expression, we first used *in situ* hybridization to detect *Sox4* mRNA. Consistent with previous reports, *Sox4* is restricted to the very base of the intestinal crypts in the stem cell zone (Fig. 1A) (18). Since the cellular resolution of *in situ* hybridization is low, we leveraged a previously characterized *Sox9*^{EGFP} reporter mouse to examine the expression of *Sox4* in different crypt cell types. Distinct *Sox9*^{EGFP} levels differentially mark active ISCs (*Sox9*^{low}), secretory progenitors/reserve ISCs (*Sox9*^{high}), TA progenitors (*Sox9*^{sublow}), Paneth cells (*Sox9*$^{EGFP\text{-}neg}$:CD24^{+}), and non-Paneth post-mitotic cells (*Sox9*$^{EGFP\text{-}neg}$:CD24^{neg}) (12, 13, 16, 22). *Sox9*^{EGFP} populations were isolated by fluorescence-activated cell sorting (FACS) and analyzed by RT-qPCR. *Sox4* was specifically upregulated in the *Sox9*^{low} population, which is consistent with active ISCs, and the *Sox9*^{high} population, which is consistent with secretory progenitor cells possessing reserve ISC function (Fig. 1B) (16, 23).

Next, we examined SOX4 protein expression by immunofluorescence. SOX4 protein was expressed in a more restricted pattern relative to mRNA. Like *Sox9*, we observed distinct expression levels of SOX4, which we characterized as “low” and “high” (Fig. 1C). SOX4^{low} cells were associated with the crypt base columnar (CBC), active ISC population, in cell positions +1-3 counting from the base of the crypt (Fig. 1C, D). Interestingly, SOX4^{high} cells were predominantly found in the supra-Paneth cell

position, which is commonly associated with early secretory progenitors, at cell positions +4-7 (Fig. 1C, D) (24). Unlike SOX9, which is expressed in villus-based EECs/tuft cells, we did not detect any *Sox4* mRNA or protein outside of the crypts (12).

Since *Sox9*^{high} cells are known to represent EEC-like secretory progenitors, we asked if the SOX4^{high} population overlapped with *Sox9*^{high} cells. However, we found no co-localization between high levels of SOX4 protein and high levels of *Sox9* mRNA, as determined by *Sox9*^{EGFP}, which is an accurate surrogate of endogenous *Sox9* mRNA and protein expression (n = 3 mice, 100 SOX4+ cells per mouse) (Fig. S1A) (12, 25). *Sox9*^{high}-expressing cells are consistent with a subset of label-retaining reserve ISCs (LRCs) located in the supra-Paneth cell position. However, not all LRCs are *Sox9*^{high} (16). We reasoned that the supra-Paneth cell location of SOX4^{high} cells suggests that they may represent the non-*Sox9*^{high} subset of LRCs. To examine this, we counted the number of SOX4^{high} cells co-localizing with LRCs, as determined by 28 days of continuous labeling with 5-ethynyl-2-deoxyuridine (EdU), followed by 10 days of washout (16, 24). Surprisingly, we found that no SOX4^{high} cells were LRCs (n = 3 mice, 100 SOX4+ cells per mouse) (Fig. S1B). We did, however, observe rare co-localization between tuft cell/reserve ISC biomarker DCKL1 and SOX4 (5.0% ± 3.0%; Fig. S1C). Finally, we addressed the possibility that SOX4^{high} cells may represent a quiescent population by co-staining with KI67, which marks all proliferating cells. 90.6 ± 3.2% of SOX4^{high} cells were found to be proliferative (Fig. S1D). Together, these data suggest that SOX4^{high} cells represent ISCs or progenitors, and are distinct from LRC *Sox9*^{high} secretory precursors. Co-localization with DCLK1 suggests that a subset of SOX4^{high} cells may represent tuft cell precursors.

Loss of Sox4 results in increased proliferation and decreased ISC function

Complete deletion of *Sox4* results in embryonic lethality (26). To enable examination of *Sox4* function in adult intestinal epithelium, we crossed *Sox4*$^{fl/fl}$ mice to mice expressing Cre recombinase driven by the *villin* promoter (*vilCre*), which is constitutively expressed in the intestinal epithelium beginning at E13.5 (27, 28). To examine *Sox4* expression, we isolated CD44- and CD44+ populations from control and Sox4cKO adult epithelium (CD326+) by FACS and confirmed CD44 enrichment by qPCR (Fig. S2A-C). CD44 marks crypt cells as well as EECs in the villi (Fig. S2A). As expected, control animals exhibited significant enrichment of *Sox4* in the CD44+ population, consistent with restricted expression of *Sox4* in the intestinal crypts (Fig. S3A). Sox4cKO animals did not express detectable levels of *Sox4* mRNA or protein (Fig. S3A, B).

To assay the role of *Sox4* in ISC/TA proliferation, we examined KI67 and EdU, which was administered as a single, 2hr pulse. Sox4cKO crypts exhibited significantly higher percentages of total proliferating cells as well as cells in S-phase (Fig. 2A, B). This increase in proliferation was reflected by increased crypt depth in Sox4cKO intestines (Fig. S3C). However, crypts isolated from Sox4cKO intestines and subjected to 3D organoid culture failed to demonstrate a significantly higher growth rate relative to controls, suggesting that hyperproliferation in the absence of *Sox4* may be masked by exogenous growth factors required for organoid culture (Fig. 2C).

Due to its known role in maintenance of stem/progenitor pools in other tissues, we sought to determine if *Sox4* regulates the size or function of the ISC pool. OLFM4, which

is an established marker of ISCs, was detected by immunofluorescence to assess ISC numbers, which were unchanged between control and Sox4cKO intestines (Fig. 2D) (29). Next, we FACS-isolated CD44+ cells from control and Sox4cKO mice and subjected them to single cell organoid forming assays to test ISC function (30). Sox4cKO cells demonstrated a significant loss of functional stemness as measured by organoid formation, producing approximately 50% fewer organoids than controls (Fig. 2E). To determine if this functional deficit was secondary to an increase in apoptosis in Sox4cKO samples, we examined cleaved-caspase-3 expression in the crypts of control and Sox4cKO intestines, as well as the number of apoptotic cells in CD44- and CD44+ populations of cells prepared for FACS. In both cases, we found no difference in apoptosis between control and Sox4cKO samples, suggesting that decreased organoid formation results from a loss of stemness and not from apoptosis (Fig. S3D & E).

To determine if organoid-forming deficit correlated with a loss of ISC gene expression, we analyzed ISC biomarkers in CD44 populations. Consistent with a loss of stemness, CD44+ cells from Sox4cKO mice expressed significantly lower levels of *Lgr5* (Fig. 2F). There was no significant change in Wnt target genes *Ascl2*, *Myc,* or *Axin2* in Sox4cKO CD44+ cells (Fig. 2F). Interestingly, Sox4cKO mice expressed significantly higher levels of *Axin2* in the CD44- population, consistent with expanded crypt size and increased proliferation (Fig. 2F) (31). Together, these data indicate that *Sox4* does not regulate ISC numbers, but is required for normal ISC function.

Sox4 regulates secretory lineage allocation

We next sought to determine if *Sox4* regulates ISC competence to differentiate

into post-mitotic lineages. Goblet cells, detected by expression of MUC2, were found at similar numbers between Sox4cKO and control intestines (Fig. 3A). CHGA$^+$ EECs were found in significantly reduced numbers in the villi of Sox4cKO mice (Fig. 3A). Numbers of DCLK1$^+$ tuft cells were significantly reduced in crypt and villus compartments of Sox4cKO intestines (Fig. 3A). Paneth cells were increased in Sox4cKO intestines (Fig. 3A), reflecting additional Paneth cells at higher positions in Sox4cKO crypts (Fig. S4A). Staining patterns for sucrose isomaltase (SIS), a brush-border enzyme associated with absorptive enterocytes, were unchanged between control and Sox4cKO samples (Fig. S4B). Additionally, villus height was unchanged, and no change was observed for *Hes1*, which specifies absorptive enterocyte fate (Fig. S3C & S4C).

To determine if secretory differentiation defects were associated with gene expression changes in known upstream regulators of differentiation, we examined sorted CD44 populations by RT-qPCR. *Atoh1,* which is a master regulator of secretory differentiation, was significantly decreased in Sox4cKO crypt cells, consistent with dysregulation in secretory lineage allocation (Fig. 3B). Transcription factors regulating goblet and Paneth cell differentiation, *Sox9* and *Gfi1* were unchanged, while *Spdef* was downregulated in Sox4cKO samples (Fig. S5). Since decreased EEC numbers were observed in Sox4cKO samples, we next examined expression of known EEC regulators *Ngn3*, *Neurod1*, and *Isl1* (32-35). While *Neurod1* exhibited a downward trend in expression in Sox4cKO samples, only *Isl1* was significantly downregulated (Fig. 3B).

In order to examine gene expression changes in Sox4cKO samples in an unbiased manner, we subjected CD44$^-$ and CD44$^+$ populations to RNA-seq. Consistent with histological and RT-qPCR observations, decreased expression of genes associated with

EEC subpopulations, including pro-EEC transcription factors *Isl1, Nkx2.2,* and *Pax6*, was the predominant change observed in Sox4cKO $CD44^+$ populations (Fig. S2D and Table S1) (34, 36). There were no significant gene expression changes in the $CD44^-$ populations consistent with the restricted expression pattern of SOX4 in $CD44^+$ intestinal crypts (Fig. S2D). Since the EEC lineage is comprised of a diverse complement of subpopulations, we re-examined Sox4cKO intestines for the presence of hormones produced by specific EEC subtypes. Compared to the modest reduction in $CHGA^+$ cell numbers, we found substantial and significant decreases in EEC subpopulations expressing GIP, CCK, SST, GCG, PYY, and GLP-2 (Fig. 3C). Interestingly, numbers of 5-HT and SUBP-expressing populations remained unchanged (Fig. 3C). Together, these data demonstrate that *Sox4* maintains ISC competence to produce proper ratios of secretory cell populations, including a number of EEC subtypes.

Notch signaling represses Sox4

ISC cell fate is regulated by extrinsic signaling pathways present in the intestinal crypts (1). As *Sox4* expression is restricted to the crypt base, we investigated if known extrinsic regulators of differentiation are capable of driving *Sox4* expression. To modulate signaling, we focused on: WNT3A, which promotes Paneth cell differentiation as well as stemness; Noggin, which suppresses BMP signaling, a negative regulator of *Sox4* in the hair follicle; and DAPT, which inhibits Notch and increases secretory cell differentiation through *Atoh1* (37-41). Intestinal organoids were established from wild type intestines for 48hrs before being treated with growth factors and inhibitors for 24hrs. During treatment, all other factors present in basal organoid cultures were removed with

the exception of EGF, which is essential for short-term organoid maintenance (42). Contrary to previous data in colon cancer cell lines and hair follicle stem cells, we found that WNT3A signaling and BMP inhibition through Noggin were insufficient in upregulating *Sox4* (Fig. 4A) (39, 43). However, Notch inhibition through DAPT treatment resulted in a 2-fold increase in *Sox4* expression (Fig. 4A). Concomitantly, we observed an increase in the number of $SOX4^{+}$ cells in DAPT-treated organoids relative to untreated controls (Fig. 4B). As expected, DAPT-treatment also produced an increase in *Ngn3* expression and decrease in *Ascl2*, consistent with the initiation of EEC differentiation and exit from an ISC state (Fig. 4A). These data are consistent with our observation of high levels of SOX4 in the supra-Paneth cell position, as Paneth cells produce Notch ligands that drive high levels of Notch signaling in ISCs (40). Together, this suggests that loss of contact with Notch ligands in the immediate ISC niche may drive upregulation of SOX4 *in vivo*.

Tuft cell hyperplasia is regulated by Sox4

Since a subset of $SOX4^{high}$ cells were found to co-localize with DCLK1 and $DCLK1^{+}$ tuft cells were decreased in Sox4cKO intestines, we examined the regulatory role of *Sox4* in tuft cell differentiation. IL-13 is known to induce tuft cell specification, and is produced by Group 2 innate lymphoid cells (ILC2) that are recruited to the intestinal epithelium in response to tuft cell secreted IL-25. This mechanism produces an extrinsic feed-forward loop that signals through unknown intrinsic response elements in ISCs and progenitors to drive tuft cell hyperplasia following parasitic helminth infection (2, 5). We treated wild type intestinal organoids with IL-13 to determine its effect on

Sox4. As previously reported, IL-13 induced expression of tuft cell-specific *Dclk1* (Fig. S6A). Compellingly, IL-13 also induced a dose-dependent upregulation of *Sox4*, driving an approximately 3-fold increase at the optimal dose (Fig. S6A). Levels of *Sox4* upregulation correlated closely with *Dclk1* induction. As expected, goblet cell *Muc2* was the only non-tuft lineage specific gene also upregulated in response to IL-13 (Fig. S6B) (2). To determine if *Sox4* is required for proper *Dclk1* upregulation in response to IL-13, we next treated Sox4cKO and control organoids with IL-13. While control organoids exhibited robust induction of *Dclk1* and the tuft cell transcription factor *Pou2f3*, Sox4cKO organoids failed to significantly upregulate these genes (Fig. 5A).

Next, we sought to determine if *Sox4*-driven tuft cell responses were physiologically necessary for clearance of helminth infection. Sox4cKO and control animals were infected with *N. brasiliensis*, a helminth known to induce tuft cell-mediated type II immune reactions (2, 4, 5). Intestines were examined 7 days post-infection (d.p.i.), at the peak of tuft cell hyperplasia, when most worms are typically cleared from wild type mice (5). We first examined $SOX4^+$ cells in control mice. Helminth infection drove a 4-fold increase in the number of $SOX4^+$ cells per crypt (Fig. 5B). As expected, $DCLK1^+$ cell numbers were largely upregulated in control animals (Fig. 5C). In contrast, Sox4cKO intestines exhibited an attenuated increase in tuft cell numbers (Fig. 5C). To determine if impaired tuft cell hyperplasia in Sox4cKO animals affected worm burden at 7d.p.i., we quantified the number of adult worms present in the intestines of infected animals. In order to match worm clearance data with tuft cell hyperplasia in the same mouse, duodenal and ileal tissues were harvested to establish worm counts and jejunal tissues were used for histological studies, discussed above. Control animals exhibited a near-

total clearance of adult worms, while Sox4cKO animals bore a significantly higher worm burden at 7d.p.i. (Fig. 5D). These data indicate that *Sox4* is a transcriptional regulator of tuft cell differentiation at physiological baseline, and plays an important role in tuft cell hyperplasia during helminth infection.

Single cell transcriptomics segregate Sox4-expressing cells and tuft cells from Atoh1-expressing cells

Early tuft cell specification remains an enduring paradox in intestinal epithelial differentiation. *Atoh1* is accepted as the master regulator of early secretory fate in the intestine, and is de-repressed following loss of Notch signaling (44). While tuft cells fail to differentiate properly in the presence of constitutively active Notch, multiple lines of evidence demonstrate that a majority of tuft cells are not derived from *Atoh1*-expressing progenitors (3, 7). These data suggest the existence of Notch-repressed, *Atoh1*-independent tuft cell regulatory pathways that remain uncharacterized. Since our data demonstrate that *Sox4* is upregulated following Notch inhibition and is involved in tuft cell differentiation, we hypothesized that it might be expressed in an *Atoh1*-independent manner. We examined SOX4 expression in *Atoh1*^{GFP} intestines, and found that 55.3 ± 6.1% of SOX4^{high} cells were *Atoh1*-negative, with the remainder co-localizing with Atoh1-GFP (Fig. 6A).

To examine *Sox4* and *Atoh1* populations more thoroughly, we reanalyzed previously published single cell RNA-seq (scRNA-seq) data from primary mouse intestinal epithelial cells (45). Cells expressing any combination of *Sox4*, *Atoh1*, and *Dclk1* were selected for analysis. As a benchmark for comparison, *Sox4-/Atoh1-/Dclk1+*

cells were considered mature tuft cells, based on lack of either *Sox4* or *Atoh1* progenitor genes. Following quality control steps to exclude cells with low transcriptomic diversity (see Methods), we analyzed 332 cells by t-SNE to determine population heterogeneity. As expected, cells were generally separated into larger populations based on *Atoh1* expression, with *Sox4-/Atoh1+* (light and dark green dots) and *Sox4+/Atoh1+* cells (purple and brown dots) clustering differentially from *Sox4+/Atoh1-* cells (orange and yellow dots) (Fig. 6B & C). Within these clusters, *Sox4+/Atoh1-/Dclk1-* cells (orange dots) associated more closely with mature tuft cells (pink dots), which were interspersed with a large number of *Sox4+/Atoh1-/Dclk1+* cells (yellow dots) (Fig. 6B). A small number of cells expressing all three genes (brown dots) also clustered with mature tuft cells. Consistent with their proposed ISC/progenitor status, *Sox4+/Atoh1-/Dclk1-* and all *Atoh1+* populations were enriched for proliferative markers *Ki67, Ccna1,* and *Ccnd1*, while mature tuft cells largely did not express these genes (Fig. 6D). To confirm the identity of putative tuft cell populations, we examined the expression of known tuft cell markers *Pou2f3*, *Plcb2*, and *Trpm5* by tSNE analysis. As expected, mature tuft cells and *Sox4+/Atoh1-/Dclk1+* cells were enriched for other tuft cell markers (Fig. 6E). We also noted a second population, enriched for both tuft cell and proliferative markers, and expressing high levels of *Sox4* (Fig. 6C, D, & E). By contrast, mature EEC markers were enriched in *Sox4-/Atoh1+* and *Sox4+/Atoh1+* populations and EEC transcription factors, including *Neurod1, Ngn3,* and *Nkx2-2* were enriched in *Sox4+/Atoh1+* and a small cluster of *Sox4+/Atoh1-* cells (Fig. 6F and Fig. S7A). Two smaller populations were identified, both of which expressed higher levels of *Atoh1* and lower levels of *Sox4* along with known Paneth cell genes, suggesting that they may represent mature Paneth cells or

Paneth progenitors (Fig. S7B). Additionally, goblet cell markers *Muc2* and *Tff3* were expressed more promiscuously across all populations examined, but were especially enriched within *Atoh1*+ subpopulations (Fig. S7C). Taken together, these data demonstrate that unsupervised transcriptomic clustering of single cells results in sub-classification of *Sox4*+ cells based on *Atoh1* expression status. Furthermore, *Sox4+/Atoh1-* cells cluster more closely with mature tuft cells than *Sox4+/Atoh1+* and *Sox4-/Atoh1+* cells, which are more enriched for non-tuft secretory markers, including those associated with EECs, Paneth cells, and goblet cells. These data support functional data from Sox4cKO in suggesting that *Sox4* may be a marker of early tuft cell progenitors and that it may act independently of *Atoh1* to drive tuft cell fate decisions.

Discussion

Here we show that genetic ablation of *Sox4* in the intestinal epithelium results in subtle but significant increases in the number of proliferating cells and Paneth cells, as well as a reduction in EEC and tuft cell lineages. Our study supports the concept that *Sox4* regulates ISC competence and influences cell fate decisions of rare secretory progenitors to generate subsets of the EEC lineage and tuft cells in the intestinal epithelium. Outside of normal physiology, *Sox4* has an essential role in pushing lineage allocation toward tuft cell differentiation in response to helminth infection.

Our group and others have reported *Sox4* expression in *Lgr5*+ CBCs, but a role for *Sox4* in ISCs has yet to be established (18, 22). We demonstrate that the numbers of ISCs quantified by OLFM4 immunostaining are unchanged between Sox4cKO and control samples. These data suggest that the associated increase in proliferating cells in

Sox4cKO crypts is likely due to increased progenitor numbers and not more robust ISC activity. Rather, *in vitro* assay of ISC function demonstrated impaired organoid-forming efficiency of single ISCs isolated from Sox4cKO mice. As this dysfunction is not associated with a change in the number of apoptotic cells between control and Sox4cKO mice, it is likely that decreased organoid-forming efficiency is secondary to a loss of stemness and premature differentiation in *Sox4* deficient ISCs. In this regard, Sox4cKO mice demonstrate reduced *Lgr5* mRNA. *Lgr5* is highly expressed in ISCs and serves as a receptor for R-Spondin ligands that potentiate Wnt-signaling, which is in turn required to maintain stemness (46). These data indicate that *Sox4* is not required for maintaining ISC numbers, but suggests that *Sox4* plays a role in maintaining ISC competence by positively regulating *Lgr5* levels. Functionally, this role would be consistent with *Sox4* knockout phenotypes in other tissues, including the developing heart, skeletal muscle, nervous, renal, and hematopoietic systems, which often exhibit loss of progenitor competence and reduced post-mitotic cell numbers (19, 20, 26, 47-49).

Sox4cKO intestines also demonstrate decreased EEC and tuft cell numbers, implying that *Sox4* plays a key role in the specification of these secretory lineages. RNA-seq and qPCR analysis demonstrate that *Sox4* regulates the expression of the EEC transcription factors, including *Isl1, Pax6,* and *Nkx2-2*. In particular, *Isl1* has been shown to have an essential role in maintaining progenitor cells in the pancreas, heart and nervous system. In the intestine, *Isl1* deletion leads to loss of rare EECs subtypes expressing GIP, CCK, SST, GCG, PYY, and GLP-2 (34). We demonstrate that loss of *Sox4* results in a marked reduction of *Isl1* expression with a concomitant reduction in the numbers of the same EEC subtypes. Together, these results indicate that *Sox4* likely

functions upstream of EEC transcription factors *Isl1, Pax6*, and *Nkx2-2* to drive EEC specification.

Sox4 has previously been identified as a Wnt-target gene, but our data suggest that WNT3A is insufficient to upregulate *Sox4* expression (50). Studies identifying *Sox4* as a target of Wnt signaling relied largely on assays that resulted in total loss of intestinal crypts, suggesting that loss of *Sox4* in these conditions was an indirect effect of crypt ablation. Our data support a model where *Sox4* is repressed by active Notch signaling. This observation is more consistent with the expression pattern of SOX4, which is confined to relatively rare cells in the crypt. Unlike Wnt target genes like SOX9, which tend to have broad crypt-based expression, Notch-repressed targets such as ATOH1, GFI1, and NGN3 are generally observed as rare, supra-Paneth position cells, consistent with our observation of $SOX4^{high}$ cells (51). Low levels of SOX4 expression in CBC ISCs, which are highly active for Notch, may suggest that this regulation is dose-dependent, or that it involves co-regulation by an unidentified factor, independent of Notch.

The mechanisms that give rise to the tuft cell lineage are largely understudied due to the relatively new characterization of tuft cells as a distinct secretory lineage. The role of intestinal tuft cells was recently described by a handful of independent studies, all of which demonstrated tuft cell hyperplasia following infection with parasitic helminths. Following infection, tuft cells "sense" the presence of parasites through chemosensory channels involving *Trmp5*, then secrete IL-25, which initiates a type II immune response to recruit IL-13 secreting ILC-2 cells (2, 4, 5). ILC-2 cells secrete IL-13, which induces tuft cell hyperplasia (2, 4, 5). *Pou2f3* is known to be essential for tuft cell differentiation,

but is also expressed by mature, villus-based tuft cells (2). Other transcription factors involved in tuft cell specification and hyperplasia remain unknown, and tuft cell progenitors have no distinct biomarkers (2). Though SOX4^{high} cells are relatively rare in homeostasis, we observed a striking increase in SOX4^{high} cell numbers following parasitic infection, coincident with the peak of tuft cell hyperplasia in control animals. We show that Sox4cKO epithelium demonstrates attenuated tuft cell hyperplasia, both in response to recombinant IL-13 *in vitro* and helminth infection *in vivo*, and that Sox4cKO mice are functionally impaired at clearing worms by 7d.p.i.

Interestingly, tuft cell differentiation is known to proceed independently of *Atoh1*, but is impaired in the setting of constitutive Notch signaling (3, 7). The possibility of alternative tuft cell differentiation pathways challenges the long-standing dogma of *Atoh1* as the master transcriptional regulator of secretory differentiation in the intestine. Our findings suggest that *Sox4* is upregulated following downregulation of Notch, and drives tuft cell specification, thereby inducing secretory differentiation independently of *Atoh1*. scRNA-seq reveals that *Sox4+/Atoh1-* and *Sox4/Atoh1+* cells represent transcriptomically distinct populations. While *Sox4+/Atoh1-* cells express genes consistent with either ISCs or tuft cells, *Sox4+/Atoh1+* cells are enriched for EEC genes, with a small but distinct subpopulation expressing genes consistent with Paneth cells. We propose a model in which *Sox4+/Atoh1-* progenitors represent tuft cell precursors, while *Sox4+/Atoh1+* progenitors differentiate to produce EEC lineages (Fig. 7). The precise regulatory relationship between *Sox4* and *Atoh1* remains to be determined, as does the processes by which early secretory progenitors are capable of responding to repressed Notch signaling to preferentially activate one gene or both. Previous characterization of

the tuft cell lineage demonstrated that a small subset of tuft cells are derived from progenitor cells that expressed *Atoh1* at some point during differentiation (7). These data are consistent with our findings that not all tuft cells are lost in Sox4cKO intestines, and that single cell transcriptomic studies identify a very small number of cells that co-express *Sox4*, *Atoh1*, and tuft cell genes. Our study supports the notion that tuft cell derivation is heterogeneous in terms of progenitor populations, but demonstrates that *Sox4* drives an *Atoh1*-independent tuft cell differentiation program that is highly active in response to parasitic infection.

References

1. Gracz AD & Magness ST (2014) Defining hierarchies of stemness in the intestine: evidence from biomarkers and regulatory pathways. *American journal of physiology. Gastrointestinal and liver physiology* 307(3):G260-273.
2. Gerbe F*, et al.* (2016) Intestinal epithelial tuft cells initiate type 2 mucosal immunity to helminth parasites. *Nature* 529(7585):226-230.
3. Gerbe F*, et al.* (2011) Distinct ATOH1 and Neurog3 requirements define tuft cells as a new secretory cell type in the intestinal epithelium. *The Journal of cell biology* 192(5):767-780.
4. Howitt MR*, et al.* (2016) Tuft cells, taste-chemosensory cells, orchestrate parasite type 2 immunity in the gut. *Science* 351(6279):1329-1333.
5. von Moltke J, Ji M, Liang HE, & Locksley RM (2016) Tuft-cell-derived IL-25 regulates an intestinal ILC2-epithelial response circuit. *Nature* 529(7585):221-225.
6. Noah TK & Shroyer NF (2013) Notch in the intestine: regulation of homeostasis and pathogenesis. *Annual review of physiology* 75:263-288.
7. Bjerknes M*, et al.* (2012) Origin of the brush cell lineage in the mouse intestinal epithelium. *Developmental biology* 362(2):194-218.
8. VanDussen KL*, et al.* (2012) Notch signaling modulates proliferation and differentiation of intestinal crypt base columnar stem cells. *Development* 139(3):488-497.
9. Westphalen CB*, et al.* (2014) Long-lived intestinal tuft cells serve as colon cancer-initiating cells. *The Journal of clinical investigation* 124(3):1283-1295.
10. Gracz AD & Magness ST (2011) Sry-box (Sox) transcription factors in gastrointestinal physiology and disease. *American journal of physiology. Gastrointestinal and liver physiology* 300(4):G503-515.
11. Sarkar A & Hochedlinger K (2013) The sox family of transcription factors: versatile regulators of stem and progenitor cell fate. *Cell stem cell* 12(1):15-30.

12. Formeister EJ, *et al.* (2009) Distinct SOX9 levels differentially mark stem/progenitor populations and enteroendocrine cells of the small intestine epithelium. *American journal of physiology. Gastrointestinal and liver physiology* 296(5):G1108-1118.
13. Gracz AD, Ramalingam S, & Magness ST (2010) Sox9 expression marks a subset of CD24-expressing small intestine epithelial stem cells that form organoids in vitro. *American journal of physiology. Gastrointestinal and liver physiology* 298(5):G590-600.
14. Bastide P, *et al.* (2007) Sox9 regulates cell proliferation and is required for Paneth cell differentiation in the intestinal epithelium. *The Journal of cell biology* 178(4):635-648.
15. Mori-Akiyama Y, *et al.* (2007) SOX9 is required for the differentiation of paneth cells in the intestinal epithelium. *Gastroenterology* 133(2):539-546.
16. Roche KC, *et al.* (2015) SOX9 Maintains Reserve Stem Cells and Preserves Radio-resistance in Mouse Small Intestine. *Gastroenterology*.
17. Lin CM, *et al.* (2013) Clinical and Prognostic Implications of Transcription Factor SOX4 in Patients with Colon Cancer. *PloS one* 8(6):e67128.
18. Munoz J, *et al.* (2012) The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent '+4' cell markers. *The EMBO journal* 31(14):3079-3091.
19. Huang J, *et al.* (2013) The transcription factor Sry-related HMG box-4 (SOX4) is required for normal renal development in vivo. *Developmental dynamics : an official publication of the American Association of Anatomists* 242(6):790-799.
20. Potzner MR, *et al.* (2010) Sequential requirement of Sox4 and Sox11 during development of the sympathetic nervous system. *Development* 137(5):775-784.
21. Wilson ME, *et al.* (2005) The HMG box transcription factor Sox4 contributes to the development of the endocrine pancreas. *Diabetes* 54(12):3402-3409.
22. Gracz AD, *et al.* (2015) A high-throughput platform for stem cell niche co-cultures and downstream gene expression analysis. *Nature cell biology* 17(3):340-349.
23. Van Landeghem L, *et al.* (2012) Activation of two distinct Sox9-EGFP-expressing intestinal stem cell populations during crypt regeneration after irradiation. *American journal of physiology. Gastrointestinal and liver physiology* 302(10):G1111-1132.
24. Buczacki SJ, *et al.* (2013) Intestinal label-retaining cells are secretory precursors expressing Lgr5. *Nature* 495(7439):65-69.
25. Ramalingam S, Daughtridge GW, Johnston MJ, Gracz AD, & Magness ST (2012) Distinct levels of Sox9 expression mark colon epithelial stem cells that form colonoids in culture. *American journal of physiology. Gastrointestinal and liver physiology* 302(1):G10-20.
26. Schilham MW, *et al.* (1996) Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. *Nature* 380(6576):711-714.
27. el Marjou F, *et al.* (2004) Tissue-specific and inducible Cre-mediated recombination in the gut epithelium. *Genesis* 39(3):186-193.
28. Penzo-Mendez A, Dy P, Pallavi B, & Lefebvre V (2007) Generation of mice harboring a Sox4 conditional null allele. *Genesis* 45(12):776-780.

29. van der Flier LG, *et al.* (2009) Transcription factor achaete scute-like 2 controls intestinal stem cell fate. *Cell* 136(5):903-912.
30. Yin X, *et al.* (2014) Niche-independent high-purity cultures of Lgr5+ intestinal stem cells and their progeny. *Nature methods* 11(1):106-112.
31. Schuijers J, *et al.* (2015) Ascl2 Acts as an R-spondin/Wnt-Responsive Switch to Control Stemness in Intestinal Crypts. *Cell stem cell* 16(2):158-170.
32. Jenny M, *et al.* (2002) Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. *The EMBO journal* 21(23):6338-6347.
33. Mutoh H, *et al.* (1997) The basic helix-loop-helix transcription factor BETA2/NeuroD is expressed in mammalian enteroendocrine cells and activates secretin gene expression. *Proceedings of the National Academy of Sciences of the United States of America* 94(8):3560-3564.
34. Terry NA, Walp ER, Lee RA, Kaestner KH, & May CL (2014) Impaired enteroendocrine development in intestinal-specific Islet1 mouse mutants causes impaired glucose homeostasis. *American journal of physiology. Gastrointestinal and liver physiology* 307(10):G979-991.
35. Ye DZ & Kaestner KH (2009) Foxa1 and Foxa2 control the differentiation of goblet and enteroendocrine L- and D-cells in mice. *Gastroenterology* 137(6):2052-2062.
36. Desai S, *et al.* (2008) Nkx2.2 regulates cell fate choice in the enteroendocrine cell lineages of the intestine. *Developmental biology* 313(1):58-66.
37. Farin HF, Van Es JH, & Clevers H (2012) Redundant sources of Wnt regulate intestinal stem cells and promote formation of Paneth cells. *Gastroenterology* 143(6):1518-1529 e1517.
38. Kazanjian A, Noah T, Brown D, Burkart J, & Shroyer NF (2010) Atonal homolog 1 is required for growth and differentiation effects of notch/gamma-secretase inhibitors on normal and cancerous intestinal epithelial cells. *Gastroenterology* 139(3):918-928, 928 e911-916.
39. Kobielak K, Stokes N, de la Cruz J, Polak L, & Fuchs E (2007) Loss of a quiescent niche but not follicle stem cells in the absence of bone morphogenetic protein signaling. *Proceedings of the National Academy of Sciences of the United States of America* 104(24):10063-10068.
40. Sato T, *et al.* (2011) Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* 469(7330):415-418.
41. van Es JH, *et al.* (2005) Notch/gamma-secretase inhibition turns proliferative cells in intestinal crypts and adenomas into goblet cells. *Nature* 435(7044):959-963.
42. Sato T, *et al.* (2009) Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* 459(7244):262-265.
43. van de Wetering M, *et al.* (2002) The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* 111(2):241-250.
44. Yang Q, Bermingham NA, Finegold MJ, & Zoghbi HY (2001) Requirement of Math1 for secretory cell lineage commitment in the mouse intestine. *Science* 294(5549):2155-2158.
45. Grun D, *et al.* (2016) De Novo Prediction of Stem Cell Identity using Single-Cell Transcriptome Data. *Cell stem cell* 19(2):266-277.

46. Carmon KS, Gong X, Lin Q, Thomas A, & Liu Q (2011) R-spondins function as ligands of the orphan receptors LGR4 and LGR5 to regulate Wnt/beta-catenin signaling. *Proceedings of the National Academy of Sciences of the United States of America* 108(28):11452-11457.
47. Bergsland M, Werme M, Malewicz M, Perlmann T, & Muhr J (2006) The establishment of neuronal properties is controlled by Sox4 and Sox11. *Genes & development* 20(24):3475-3486.
48. Jang SM*, et al.* (2013) Sox4-mediated caldesmon expression facilitates differentiation of skeletal myoblasts. *Journal of cell science* 126(Pt 22):5178-5188.
49. Sun B*, et al.* (2013) Sox4 is required for the survival of pro-B cells. *Journal of immunology* 190(5):2080-2089.
50. Van der Flier LG*, et al.* (2007) The Intestinal Wnt/TCF Signature. *Gastroenterology* 132(2):628-632.
51. Shroyer NF, Wallis D, Venken KJ, Bellen HJ, & Zoghbi HY (2005) Gfi1 functions downstream of Math1 to control intestinal secretory cell subtype allocation and differentiation. *Genes & development* 19(20):2412-2417.

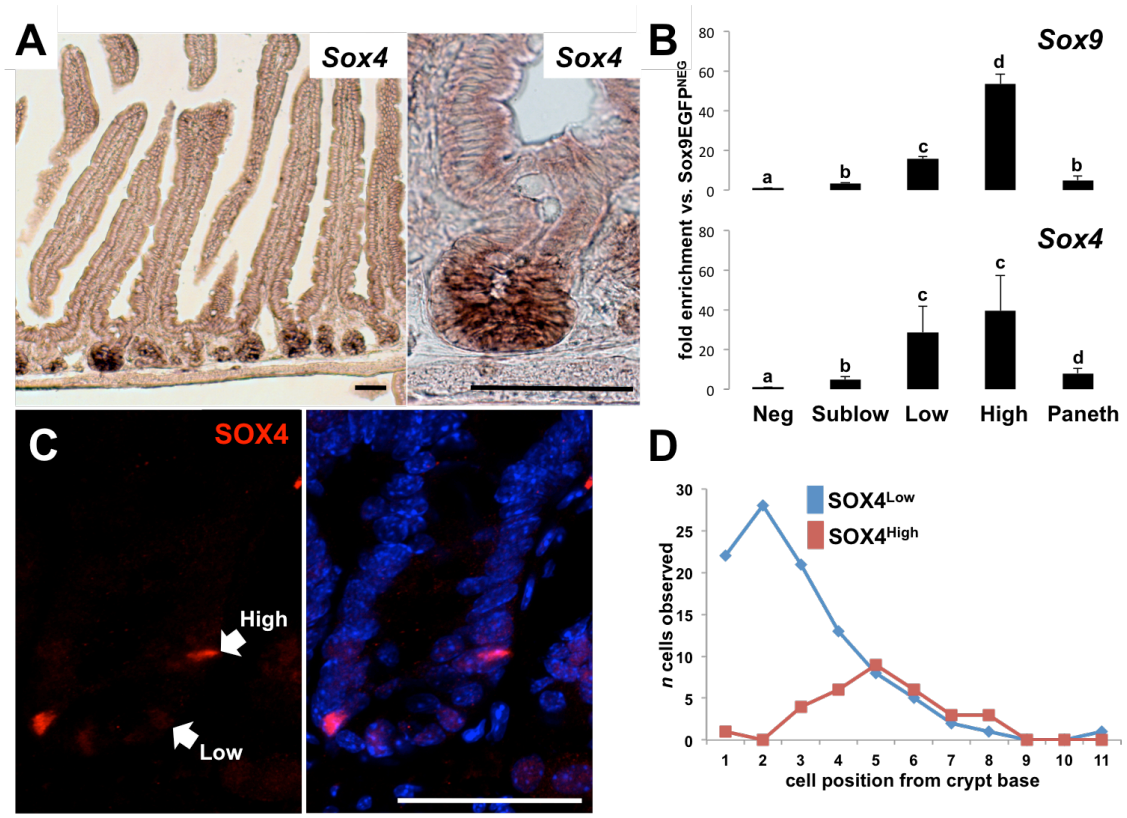

Fig. 1. *Sox4 is expressed in the ISC and early progenitor zone.* (A) *In situ* hybridization localizes *Sox4* mRNA to the base of small intestinal crypts. (B) FACS isolation of *Sox9*^{EGFP} populations demonstrates enrichment of *Sox4* in active ISCs (*Sox9*^{low}) and secretory progenitors/reserve ISCs (*Sox9*^{high}) (different letters indicate statistically significant differences, $p < 0.05$). (C) SOX4 protein is expressed in the ISC and supra-Paneth cell zones at variable levels, with (D) SOX4^{low} cells primarily localized to the CBC position (1-3) and SOX4^{high} cells occurring mainly in the supra-Paneth cell zone (4+).

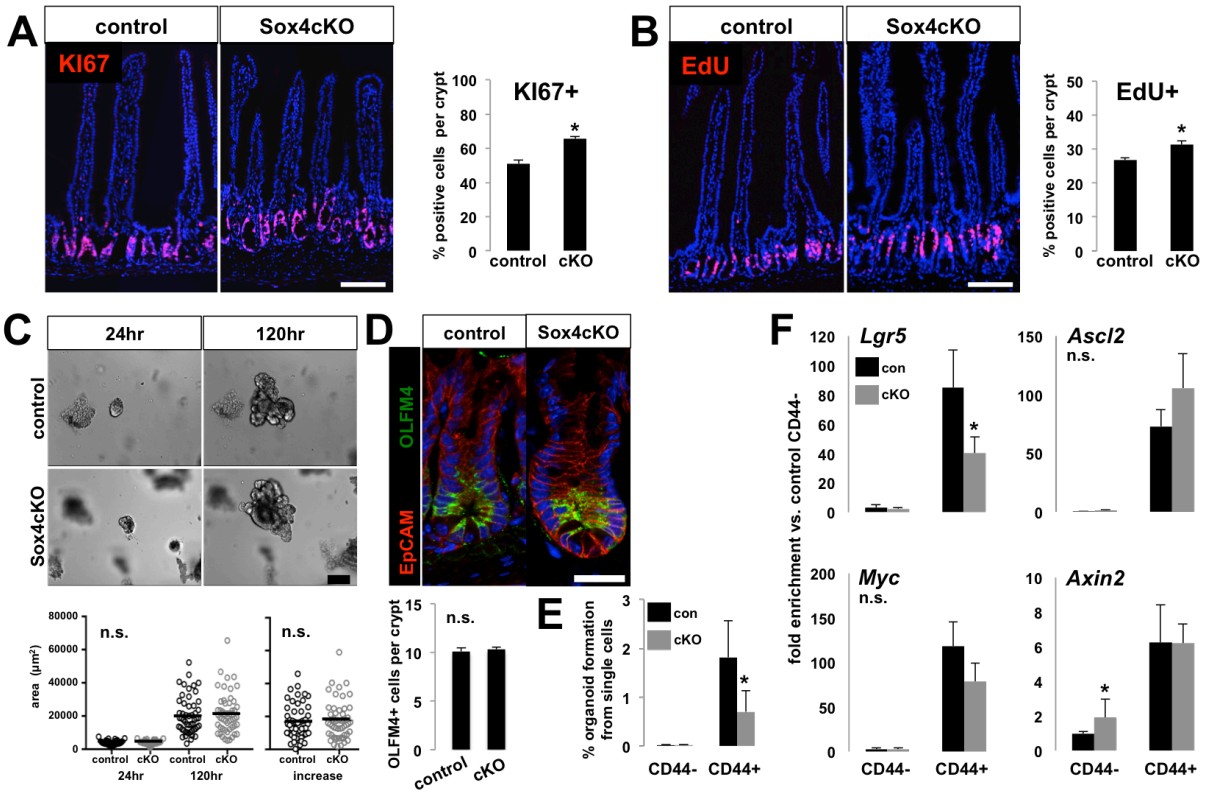

Fig. 2. *Loss of Sox4 results in increased proliferation and decrease in functional stemness, with no change in ISC numbers*. (A) Sox4cKO intestines exhibit increased total proliferation by KI67 staining, as well as (B) increased numbers of cells in S-phase, as indicated by EdU (scale bar represents 100um). (C) Whole crypts isolated from Sox4cKO intestines form organoids that undergo similar growth rates to control crypts (scale bar represents 100um). (D) Numbers of OLFM4+ ISCs remain unchanged in Sox4cKO crypts, while (E) functional stemness, as assayed by organoid forming ability, is significantly decreased in CD44+ populations isolated from Sox4cKO intestines (scale bar represents 25um). (F) *Lgr5* is downregulated in CD44+ populations from Sox4cKO intestines, while *Ascl2* remains unchanged. The Wnt target gene *Axin2* is significantly upregulated in CD44- cells from Sox4cKO intestines, consistent with increased proliferation outside of the ISC/progenitor zone. (asterisks indicate significance; $p < 0.05$).

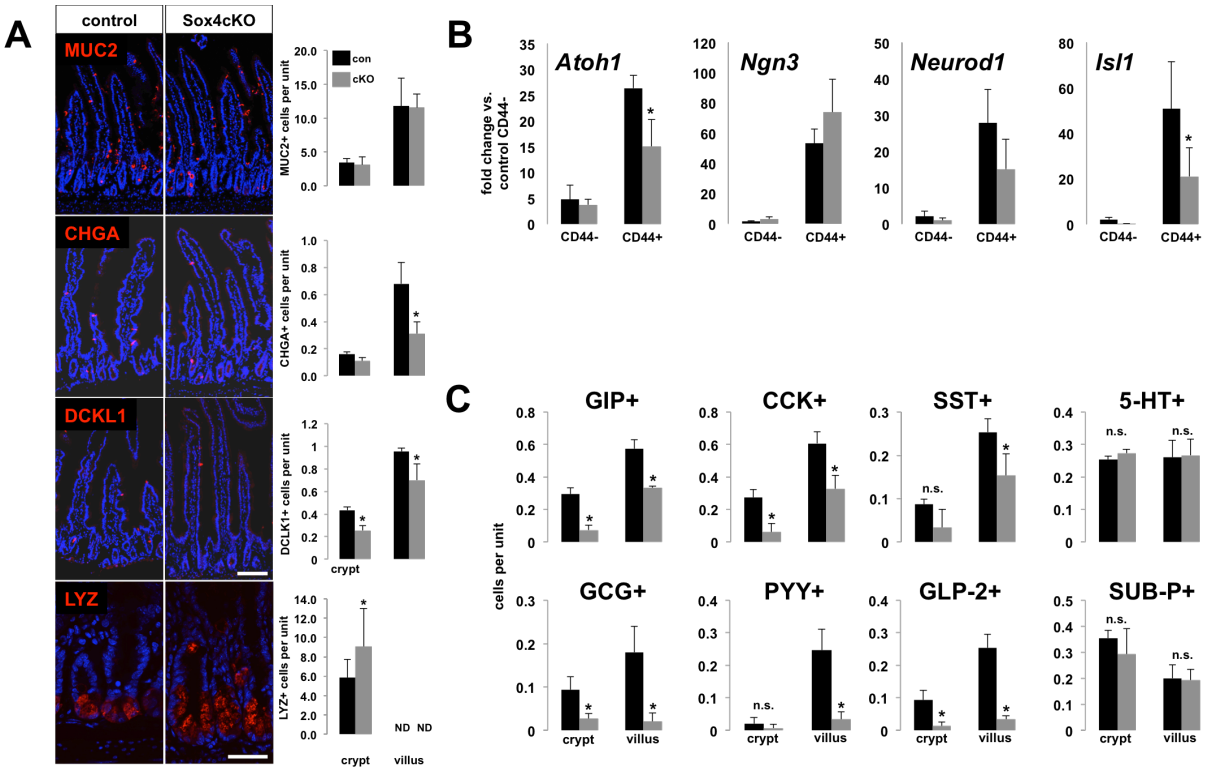

Fig. 3. *Sox4 is required for proper secretory lineage allocation.* (A) MUC2+ goblet cells remain unchanged in Sox4cKO intestines, while CHGA+ EECs are found at decreased numbers in Sox4cKO villi, and DCLK1+ tuft cells are decreased in both crypt and villus compartments of Sox4cKO intestines (scale bar represents 100um). Paneth cell numbers are increased in Sox4cKO crypts (scale bar represents 50um). (B) Gene expression analysis reveals that the master secretory lineage regulator *Atoh1* is significantly downregulated in CD44+ crypt-enriched cells in Sox4cKO. While early EEC transcription factors remain unchanged in the absence of *Sox4*, *Isl1* is significantly decreased in CD44+ Sox4cKO epithelium. (C) Assessment of EEC differentiation by immunofluorescence against markers associated with specific EEC subtypes reveals that *Sox4* is required for proper specification of a subset of EEC lineages. Conversely, numbers of cells positive for serotonin (5-HT) and Substance P (SUB-P) are unaffected in Sox4cKO intestines. (asterisks indicate significance; $p < 0.05$)

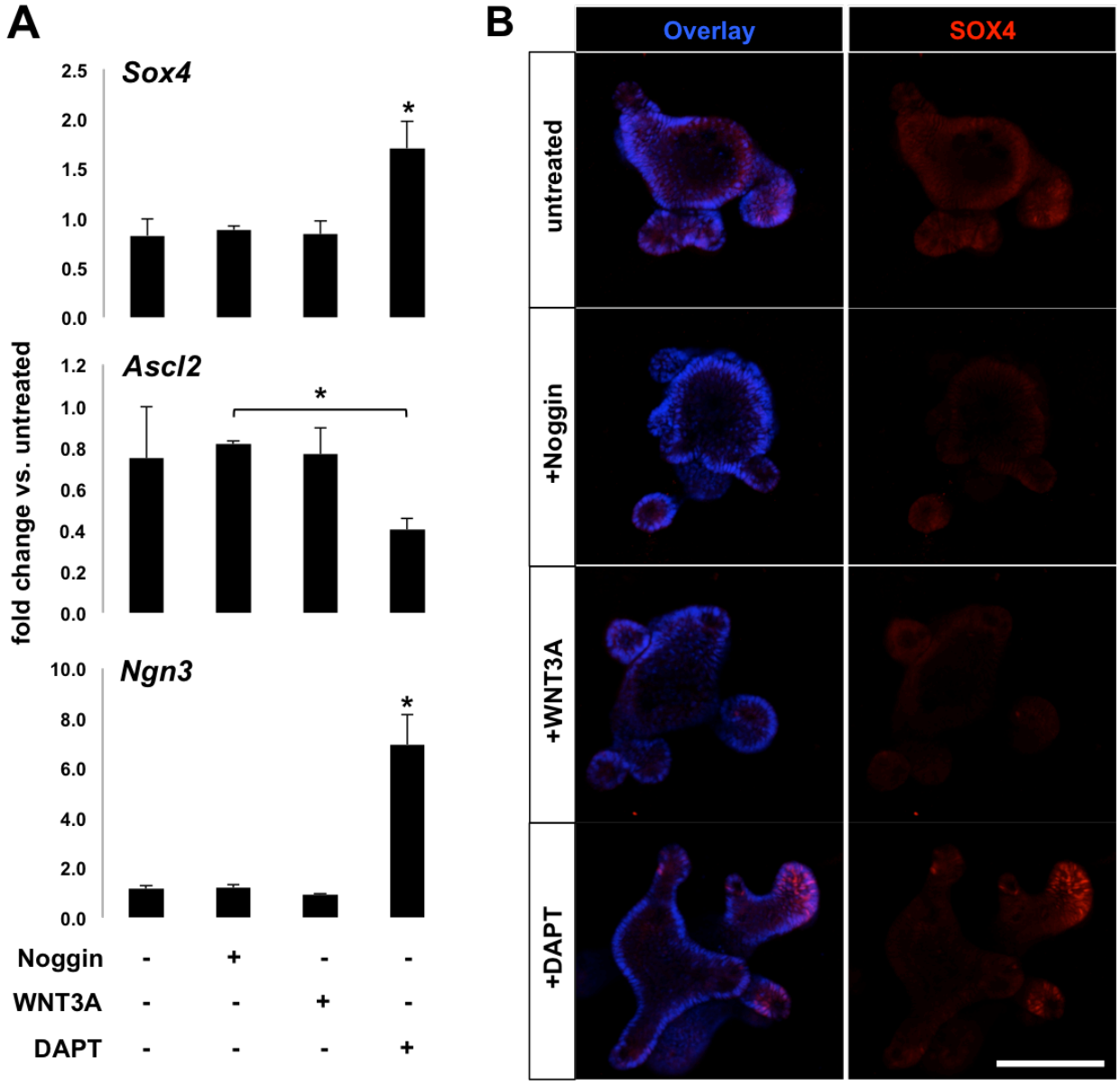

Fig. 4. *Sox4 is inhibited by Notch signaling.* (A) *Sox4* is upregulated in intestinal organoids in response to Notch inhibition with DAPT, but not in response to WNT3A or the Bmp antagonist Noggin. DAPT treatment also results in decreased expression of *Ascl2* and increased expression of *Ngn3*, consistent with loss of stemness and secretory differentiation. (B) SOX4 protein is increased in the "crypt buds" of intestinal organoids following Notch inhibition (asterisks indicate significance; $p < 0.05$; scale bar represents 100um).

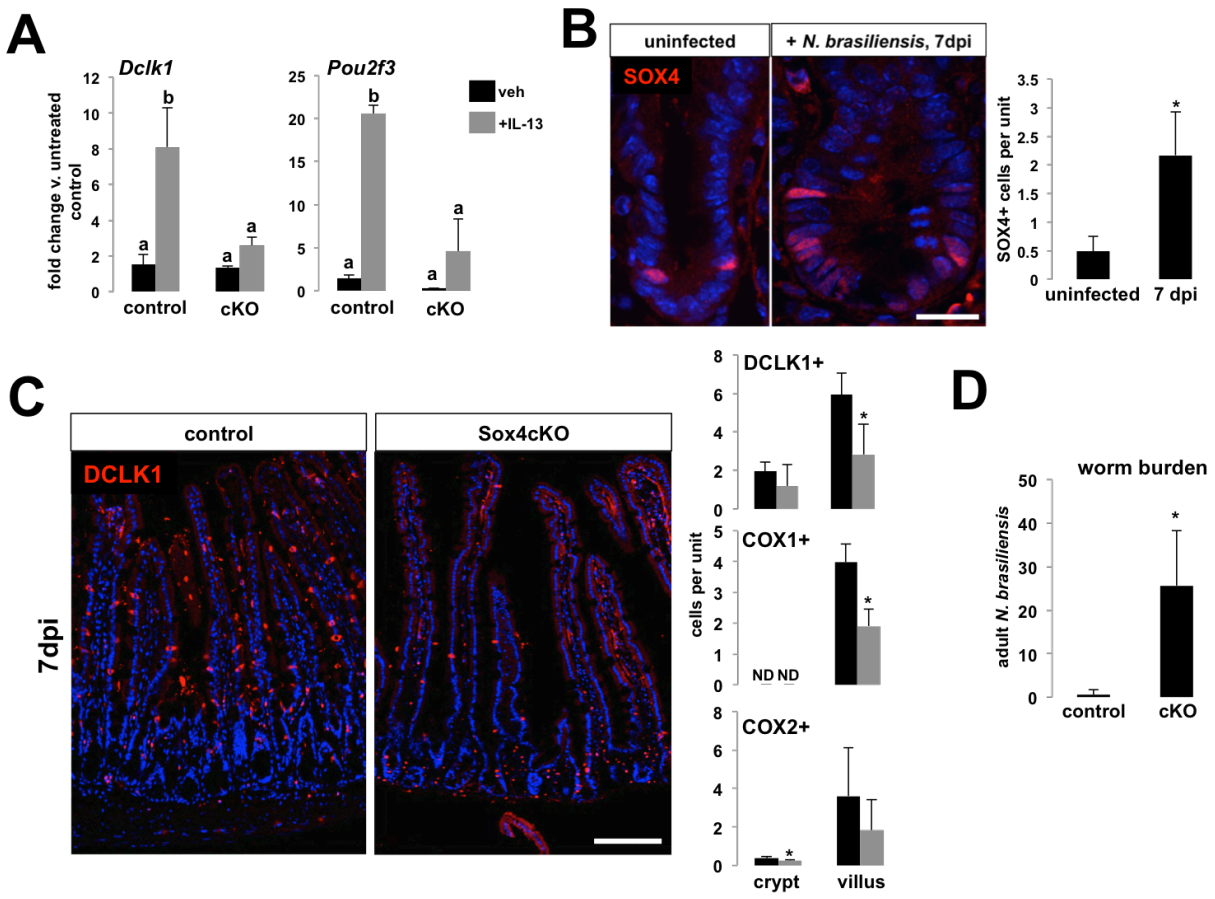

Fig. 5. *Sox4 regulates tuft cell differentiation and is required for tuft cell hyperplasia following helminth infection.* (A) Tuft cell markers *Dclk1* and *Pou2f3* are upregulated in control, but not Sox4cKO organoids, in response to treatment with recombinant IL-13 (different letters indicate statistically significant differences, $p < 0.05$). (B) 7 days post-infection with *N. brasiliensis*, control intestinal crypts exhibit a significant expansion of $SOX4^{high}$ cells (scale bar represents 20um). (C) Control mice demonstrate expected tuft cell hyperplasia at 7dpi, where Sox4cKO intestines have significantly fewer tuft cells as quantified by DCLK1, COX1, and COX2 (ND = no data; COX1+ cells are not found in intestinal crypts. Scale bar represents 100um). (D) Sox4cKO mice also fail to fully clear adult worms by 7dpi (asterisks indicate significance, $p < 0.05$, n=3 mice per group).

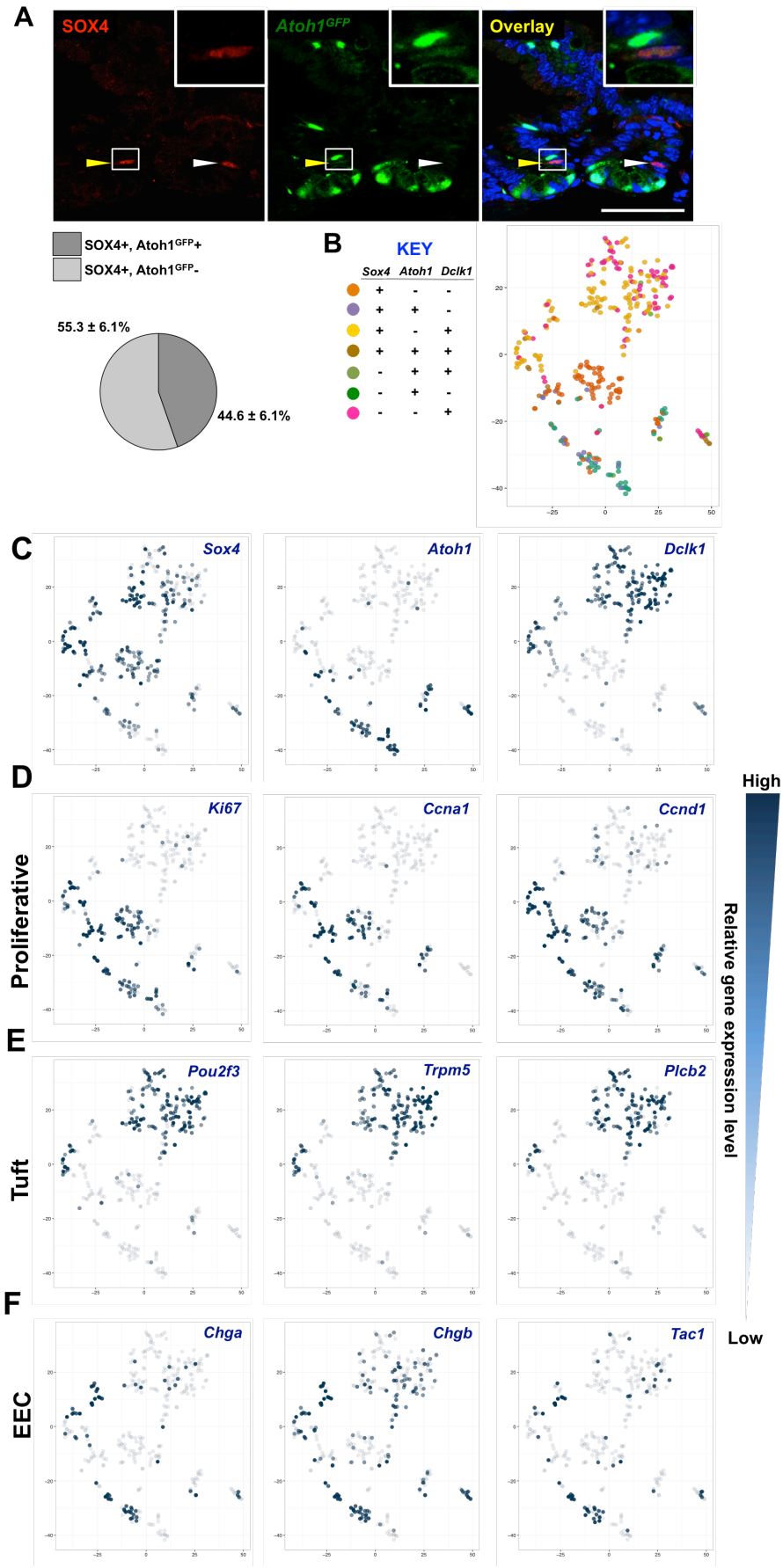
A
SOX4
Atoh1GFP
Overlay
SOX4+, Atoh1GFP+
SOX4+, Atoh1GFP-
55.3 ± 6.1%
44.6 ± 6.1%
B
KEY
Sox4 Atoh1 Dclk1
C
Sox4
Atoh1
Dclk1
D
Proliferative
Ki67
Ccna1
Ccnd1
E
Tuft
Pou2f3
Trpm5
Plcb2
F
EEC
Chga
Chgb
Tac1
High
Low
Relative gene expression level

Fig. 6. *Atoh1+ and Sox4+ populations represent distinct secretory progenitors.* (A) Co-detection of SOX4 and *Atoh1*, using a transgenic *Atoh1*^{GFP} allele, reveals that 55.3 ± 6.1% of SOX4+ cells co-express *Atoh1* (yellow arrowhead indicates SOX4+/Atoh1+ cell; white arrowhead indicates SOX4+/Atoh1- cell. Scale bar represents 25um). (B) scRNA-seq data demonstrate differential clustering of *Sox4+/Atoh1-* and *Sox4+/Atoh1+* populations based on single cell transcriptomes. (C) Distribution of *Sox4* and *Atoh1* expression correlates with classification of single cells. (D) Proliferative markers *Ki67* and *Axin2* are not associated with cells consistent with mature tuft cells as indicated by (E) expression of *Dclk1*, *Plcb2*, *Pou2f3*, and *Trpm5*, but associate with a subpopulation of cells also expressing *Sox4* and *Atoh1*. (F) EEC markers are associated mainly with *Sox4-/Atoh1+* cells, while some early EEC transcription factors, including *Neurod1* and *Nkx2-2*, also associate with a small cluster of *Sox4+/Atoh1-* cells.

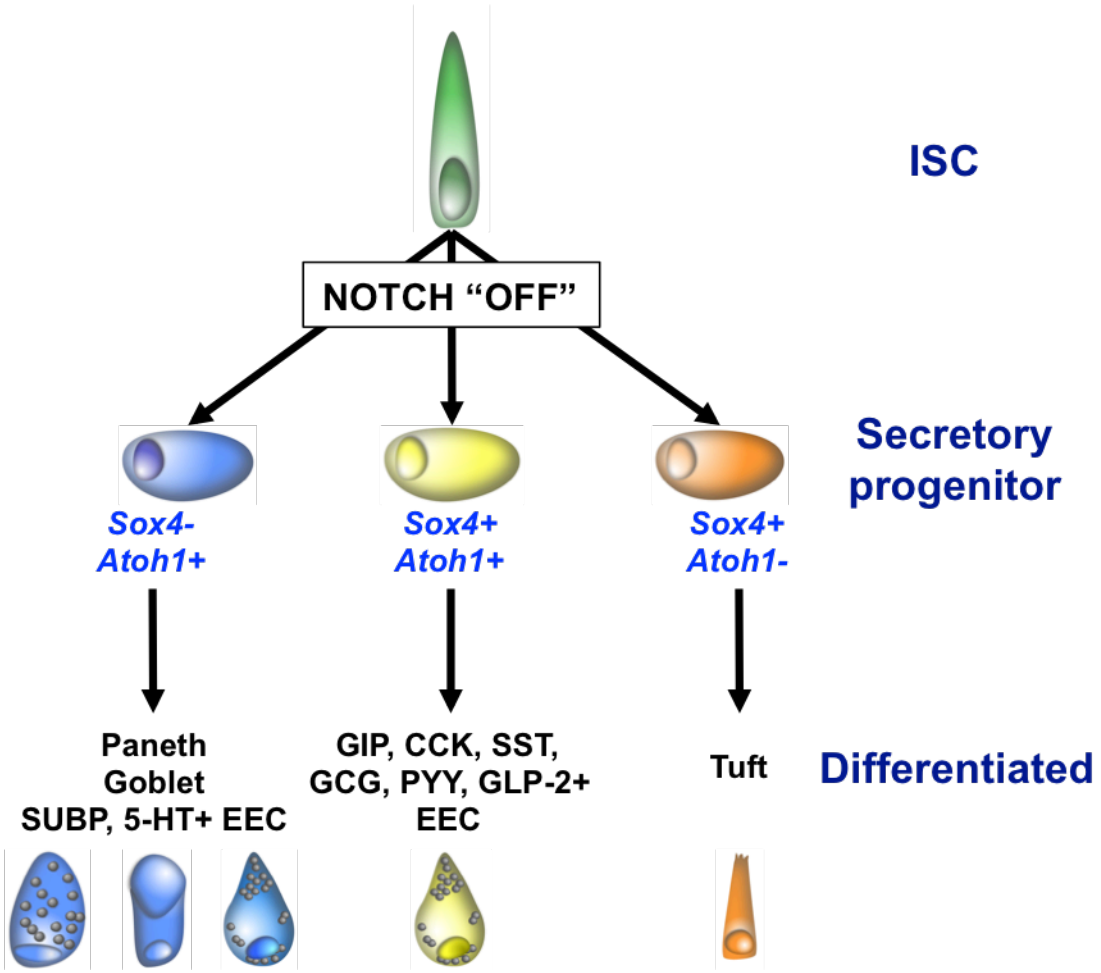

Fig. 7. *Model for differential specification of Sox4+ progenitors.* We proposed a model where secretory differentiation, initiated by loss of Notch signaling, leads to (1) upregulation of *Atoh1* without *Sox4*, resulting in Paneth, goblet, or EEC (including 5-HT and SUB-P+ cells) differentiation, (2) upregulation of *Atoh1* and *Sox4*, resulting in EEC (including GIP, CCK, SST, GCG, PYY, and GLP-2+ cells) differentiation, or (3) upregulation of *Sox4* without *Atoh1,* resulting in tuft cell differentiation.